

Title	Efficient delivery of scalable video using a streaming class model
Authors	Quinlan, Jason J.;Zahran, Ahmed H.;Sreenan, Cormac J.
Publication date	2018-03-08
Original Citation	Quinlan, J., Zahran, A. and Sreenan, C. (2018) 'Efficient Delivery of Scalable Video Using a Streaming Class Model', Information, 9(3), 59 (56pp). doi: 10.3390/info9030059
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://www.mdpi.com/2078-2489/9/3/59 - 10.3390/info9030059
Rights	© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). - http://creativecommons.org/licenses/by/4.0/
Download date	2023-05-04 21:39:30
Item downloaded from	http://hdl.handle.net/10468/5587





UCC

University College Cork, Ireland
 Coláiste na hOllscoile Corcaigh

Article

Efficient Delivery of Scalable Video Using a Streaming Class Model

Jason J. Quinlan ^{*} , Ahmed H. Zahran  and Cormac J. Sreenan

Department of Computer Science, University College Cork, T12 YT57 Cork, Ireland; a.zahran@cs.ucc.ie (A.H.Z.); cjs@cs.ucc.ie (C.J.S.)

* Correspondence: j.quinlan@cs.ucc.ie; Tel.: +353-21-4205384

Received: 7 February 2018; Accepted: 6 March 2018; Published: 8 March 2018

Abstract: When we couple the rise in video streaming with the growing number of portable devices (smart phones, tablets, laptops), we see an ever-increasing demand for high-definition video online while on the move. Wireless networks are inherently characterised by restricted shared bandwidth and relatively high error loss rates, thus presenting a challenge for the efficient delivery of high quality video. Additionally, mobile devices can support/demand a range of video resolutions and qualities. This demand for mobile streaming highlights the need for adaptive video streaming schemes that can adjust to available bandwidth and heterogeneity, and can provide a graceful changes in video quality, all while respecting viewing satisfaction. In this context, the use of well-known scalable/layered media streaming techniques, commonly known as scalable video coding (SVC), is an attractive solution. SVC encodes a number of video quality levels within a single media stream. This has been shown to be an especially effective and efficient solution, but it fares badly in the presence of datagram losses. While multiple description coding (MDC) can reduce the effects of packet loss on scalable video delivery, the increased delivery cost is counterproductive for constrained networks. This situation is accentuated in cases where only the lower quality level is required. In this paper, we assess these issues and propose a new approach called Streaming Classes (SC) through which we can define a key set of quality levels, each of which can be delivered in a self-contained manner. This facilitates efficient delivery, yielding reduced transmission byte-cost for devices requiring lower quality, relative to MDC and Adaptive Layer Distribution (ALD) (42% and 76% respective reduction for layer 2), while also maintaining high levels of consistent quality. We also illustrate how selective packetisation technique can further reduce the effects of packet loss on viewable quality by leveraging the increase in the number of frames per group of pictures (GOP), while offering a means of reducing overall error correction and by providing equality of data in every packet transmitted per GOP.

Keywords: scalable video; lossy networks; layered coding; error resilience; layer distribution; rateless coding; streaming classes

1. Introduction

Recent years have featured a dramatic rise in the volume of video streaming traffic over the Internet and mobile networks. This increase contributes to a widely acknowledged bandwidth “crunch” at the network edge and is enabled by new devices that feature a large diversity in their capabilities. However, the increase escalates many transmission issues faced by media streaming applications. The current model of transmitting multiple versions of the same video to different devices is overburdening the transmission network and is causing data to be lost during transmission. The effects of this duplication of data is being viewed in the reduction of achievable quality for each of the received streams. Hence, using adaptive video streaming schemes [1] that can adjust, or scale, the achievable quality of the media stream to the available bandwidth evolves as a crucial need for both

transmission networks and streaming applications. Multi-bitrate streaming (*adapting*) and layered coding (*scaling*) are the two primary streaming models that support video scalability. The following provides a brief overview of each technique:

- i. *Multi-bitrate (MBR) streaming* is a mechanism by which a media clip is encoded as several streams each with a different bitrate and a distinct quality version of the original media clip, sometimes known as Simulcast. Simulcast permits the Server to simultaneously stream multiple versions, or encodings, of the same video file. Early adoption of this mechanism limited the video choice of the user to only one of the available bit-rates. Subsequent change in quality required user video re-selection. A highly efficient [2] and widely used implementation of this concept is the H.264/MPEG-4 Part 10 or AVC (*Advanced Video Coding*) compression standard [3,4], while the standardised Dynamic Adaptive Streaming over HTTP (DASH) provides a clearly defined packetisation mechanism for stream delivery [5,6].
- ii. *Layered coding* [7] provides a means of adapting stream quality by adjusting the stream bitrate. Layered coding provides a means of encoding numerous fidelity (quality) levels as one stream. In layered coding, a high quality media clip is fragmented into N layers, which consist of a single base layer and $N - 1$ enhancement, or enhanced, layers. The base layer generally supports coarse minimal quality. The reception of the subsequent enhancement layers increase the viewable quality by providing an increase in temporal, spatial or quality dimensionality. Thus, stream quality adaption of layered coding is provided by means of layer selection.

A major difference between MBR and Layered Coding is the mechanism for stream adaptation. Both mechanisms communicate with the Server for the initial stream quality level. With MBR, a feedback mechanism to the Server is required for selection of the next quality level. The quality of the stream will only adapt once a request for the next segment is sent to the Server and successfully received by the Client, while, for Layered Coding, the composition of the stream permits adaptation of the stream in real time on the Client-side. Adaptation of stream quality can be achieved by dropping, or deleting, specific layers within the bitstream.

In this paper, we investigate the transmission of well-known scalable video models over a lossy network and determine how the variation in viewable quality is affected by packet loss. The overall approach is focused on the strategic packetisation of the underlying scalable video and how to best utilise error resiliency to maximise viewable quality. In particular, we examine the manner in which scalable video is packetised for transmission over lossy networks and highlight new techniques that reduce the impact of packet loss on scalable video by selectively choosing how to packetise the data and which data to transmit. We also exploit redundancy techniques, such as error resiliency, to enhance the stream quality by ensuring a smooth play-out with fewer changes in achievable video quality. We reduce delivery cost and improve overall viewable quality through the creation of novel segmentation and encapsulation techniques that increase the viewable quality of existing scalable models by fragmenting and re-allocating the video sub-streams based on user requirements, available bandwidth and variations in loss rates. We offer novel packetisation techniques that reduce the effects of packet loss on viewable quality by leveraging the increase in the number of frames per group of pictures (GOP) and by providing equality of data in every packet transmitted per GOP. These provide novel mechanisms for packetising and error resiliency, as well as providing new applications for existing techniques such as Interleaving and Priority Encoded Transmission.

2. Background and Related Work

We begin this section by providing a brief overview of the scalable/layered models in the literature and the typical effects of data/packet loss on the achievable quality of each model.

2.1. Scalable Video Coding (SVC)

Scalable Video Coding (SVC) is an extension to the H.264/MPEG-4 Part 10 or AVC (Advanced Video Coding) compression standard. The SVC standard [8,9] provides a mechanism for devices to adjust stream quality by varying the bitrate of the media stream so as to suit network conditions and device requirements. Similar to AVC, SVC only standardises the bit-stream format [10] and the central decoding process, thus permitting transmission protocols to be designed independently, such as a variant of RTP for SVC [11,12]. Ref. [13] provides information on the transport structure of SVC using RTP. Benefits of SVC include permitting devices to utilise pre-buffered lower layer data when requesting an increase in stream quality. An example of this benefit is where the base layer has been received and there exists sufficient bandwidth and time to increase viewable quality by receiving an additional higher layer. Only the additional layer needs to be transmitted and not the base layer, thus minimising the bandwidth transmission cost, thus providing the benefit of cumulative stream transmission, where different layers can be combined to increase overall viewable quality. This also allows devices with differing stream requirements to selectively choose between the layers on offer, in order to maximise their respective stream quality without requesting additional data to be transmitted. As well as implementing the key concepts of layered coding, SVC also inherits the GOP functionality of AVC. Looking to the near future, a scalable extension to High Efficiency Video Coding (HEVC) is posed for release and further development [14–17].

The goal of SVC is to provide prioritised inter-dependent layers that mandate graceful degradation of viewable quality during periods of network loss, i.e., as the percentage loss rate increases, the viewable quality of SVC is incrementally reduced due to the lower available bit-rate. In reality, this is not so. SVC is acutely affected by stream quality degradation, as the percentage of datagram loss increases. This is due to the layer dependency inherent in SVC, where the loss of a lower layer adversely affects the decodable quality, as the higher quality layer that depends upon it is unable to extract frame data and as such is un-decodable. In a mobile context, this is a significant factor that would affect its adoption. Table 1 presents the notation used in this section. Figure 1a highlights a four-layer example of SVC. The interdependency of the individual layers are determined by the selection of stream scalability of the original encoding, i.e., temporal, spatial and quality scalability. Typically, there are only two dependency scenarios, either:

1. Every layer is dependent on all lower layers, i.e., to decode any specific higher layer would require all lower layers. This dependency would occur if only one of the stream scalability options is utilised.
2. Only a subset of lower layers are required to decode a higher layer, as occurs when more than one the stream scalability options is used.

Table 1. Notation.

N	The number of Scalable Video Coding (SVC) layers per Group of Pictures (GOP)
$L_{l,x}$	Transmission cost of SVC Layer l , L_l , for frame x
<i>Section</i>	A segment or a reduced piece of an SVC Layer
$S_{l,x}$	Byte-size of a Layer section of SVC Layer l for frame x
l	Integer value corresponding to the layer number of L_l
<i>GOP</i>	The number of frames per GOP
D_c	A complete description, containing sections from layers 1 to N
q	Number of SVC layers required to decode Layer q

In addition, irrespective of encoding scalability, the base layer, L1 is required to decode all higher layers, and the highest layer index, Layer 4, from our example, requires all lower layers to decode.

If we assume that quality layer q is dependent on all lower layer $1 \dots q - 1$, then the transmission cost for SVC to decode quality layer q can be seen as

$$SVC(q) = \sum_{frame=1}^{GOP} \left(\sum_{l=1}^q L_{l,frame} \right) \quad (1)$$

or using a simple single frame per GOP example

$$SVC(q) = \sum_{l=1}^q L_l. \quad (2)$$

Thus, to recap, SVC provides:

1. Prioritised inter-dependent Layers. The viewable quality of higher (enhancement) layers can be dependent on the successful receipt of lower layers.
2. Non-graceful degradation of viewable quality. Dependent on the loss rates for each layer per frame, the variation in viewable quality can be large.
3. Viewable quality not reflective of loss rate. As the viewable quality is dependent on the maximum viewable layer decodable and loss may affect individual layer per frame differently, the probability of viewable quality being reflective of the loss rate is low.
4. A means for efficient bandwidth utilisation in networks. The greatest application of SVC is to reduce transmission cost when heterogeneous devices request live streaming such as for concerts, sporting events and TV.
5. A mechanism for the stream quality of a single user to adapt quickly in the presence of transmission loss, such that the content of the media is consistently viewable, even though the quality is reduced.

As we have seen, SVC is acutely affected by stream quality degradation due to packet loss. Prior to introducing the next streaming model, Multiple Description Coding (MDC), which specifically focuses on overcoming the impact of packet losses without having to resort to retransmissions. We first consider mechanisms that reduce the effects of packet loss.

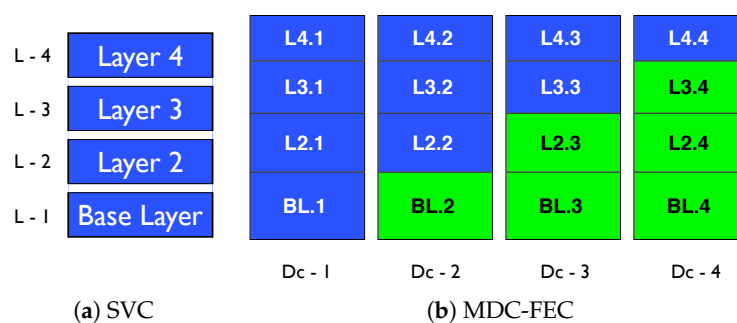


Figure 1. An example of a four-layered SVC stream encoded as MDC-FEC (blue/dark colour denotes original SVC data, green/light colour denotes additional FEC data).

2.2. Error Correction

Generally, transmission errors encountered by packet data are handled by two mechanisms: Forward Error Correction (FEC) [18] and automatic repeat request (ARQ). We shall introduce FEC later in this section. Transmission control protocol (TCP) is a key transport protocol that implements an ARQ scheme to achieve reliability. In [19], Wang et al. reveal that consistent media stream quality requires a TCP throughput twice the average media bit-rate. Additionally, the reliability and flow control mechanisms of TCP can further hinder delay sensitive real-time data [20]. These issues represent

serious limiting factors when the user has constrained bandwidth and lossy links, as it is the case for mobile video. Hence, schemes adopting FEC, such as description-based encoding, are a good alternative for media transmission over lossy links or where it is desirable to minimise latency. It is important to highlight that not all of this datagram loss is produced from non-delivery. The loss is a mixture of: late packet arrival, bit corruption in the delivered datagrams, out of order delivery and failings in prevalent transmission mechanisms, i.e., the TCP protocol is subject to high levels of loss, 21% [21] or more, where re-ordered packets are displaced by more than two positions, due to the embedded mechanisms that control the transmission window (*triple-ACK*) [22].

Several concepts have been offered to reduce the level of datagram loss, such as:

1. *Proactive datagram dropping*—Permitting the network or edge routers to determine which datagrams are of least importance, i.e., highest layers in adaptive streaming, B frames in media streaming, so as to achieve transmission energy savings and reduce packet delay [23]. The concept assumes that a reduction in data streamed will reduce transmission cost and increase delivery rate.
2. *Prioritising NAL selection in adaptive media streaming*—Sending the optimally selected subset of Network Abstraction Layer (NAL) units, so as to reduce packet transmission size [24], i.e., by reducing the quantity of packets being sent to the device. Rather than sending all NALs in a stream, the mechanism selects a subset that is beneficial, but not optimal, for the current streaming devices, which is similar to layer selection but provides more granularity in requested data.
3. *Signalling in the Network* [25]—proposed that a dedicated network node that contains sufficient knowledge of the stream packetisation mechanism to provide real-time adaptation of the stream in times of network congestion, can reduce transmission cost and subsequent levels of datagram loss, while [26] proposes an adaptive delivery mechanism based on radio resource measurements in 802.11 wireless networks, which reduces loss rates and increase user perceived quality. Both of these proposals can now benefit from the control and data separation control offered by Software Defined Networking (SDN), and its impact on routing behaviour and the underlying Quality of Service [27–29].
4. *Unequal Error Protection (UEP)*—Contrasting to the previous mechanisms, some UEP concepts add additional data to the stream, so as to reduce the datagram loss at the device, these include:
 - A. *Approximate Communication* [30]—offers a mechanism that exploits the problems inherent with corrupted data. It formulates that, when a data symbol is received, it is still a good “approximation” of the original symbol, such that by selectively altering the positions of the most significant bit (MSB) and least significant bit (LSB) to more protected positions, the confidence in accurately decoding the symbol increases.
 - B. *Partial packet recovery* [31]—offers a concept that attempts to reduce the quantity of retransmissions due to datagram corruption, also known as bit-errors, by only transmitting the portion of the packet that is corrupt. Partial packet recovery incorporates an expanded physical layer interface so as to increase confidence in determining the correct portions to request and a post-amble, located at the end of the datagram, which replicates the datagram preamble, so to be able to recover from transmission corruption in the preamble.
 - C. *Forward Error Correction*—(FEC) [32] is an example of UEP widely used by the next streaming model we shall introduce, Multiple Description Coding (MDC). Full details on FEC shall be provided in the next section.

2.3. Description-Based Streaming Models

To overcome the impact of packet losses without having to resort to retransmissions, Multiple Description Coding (MDC) [33–36] has been proposed. The key idea of MDC is introducing redundancy to the transmitted video to compensate for packet losses. MDC partitions the original N SVC layers into M descriptions [37], where the receipt of any single description provides a coarse quality representation

of the stream, i.e., base layer quality. Similar to SVC, all descriptions are required for maximum stream quality. In this regard, MDC provides a high level of consistency to stream quality by providing a mechanism, which mitigates network transmission issues albeit at a higher transmission cost in comparison to SVC.

Several variations of the MDC concept have been offered in the research literature [38] and four of the pertinent implementations are Sub-Sample, Quantisation, Transform and Forward Error Correction. We will primarily focus on Forward Error Correction (FEC) [39] as it provides a means of dynamic adaptive stream encoding, low computational complexity, the large quantity of descriptions necessary for the substantial numbers of heterogeneous media streaming devices indicated for future deployment, and it has attracted considerable attention in the literature [40–43]. Ref. [44] is one of the initial papers offering the benefits of MDC-FEC for resilient adaptive streaming.

Forward Error Correction (FEC) provides a level of error resilience proportional to layer priority. The objective of FEC is to increase the amount of transmitted data, so as to strengthen the likelihood that a subset of the data will arrive at the device, thus improving the probability that the original data can be decoded [45]. FEC is achieved by taking k existing data symbols, increasing this to n FEC data symbols (consisting of existing and redundant data symbols), such that receiving $k + 1$ FEC data symbols at the device, facilitates the recovery of the original k data symbols, as is constant with the Reed–Solomon block code rate [46]. The version of MDC associated with FEC creates a description composed of a *section* from each of the N SVC layers. A section is created by dividing a layer based on its priority, typically with the divisor from 1 to N . To provide the incremental increase in viewable quality provided by each additional description, it is typical that M equals N . Thus, the cumulative receipt of additional descriptions increases stream quality proportionally.

Figure 1b illustrates how the original four SVC layers from Figure 1a are partitioned over four description. As each original layer can be seen, i , is divided by its layer index, i , and distributed over i descriptions. Thus, for layer i , L_i , i sections are created, i.e., for layer 2, L_2 , 2 sections are created and distributed over two descriptions, while, for layer 4, L_4 , 4 sections are created and distributed over four descriptions. Each description is thus composed of layer sections. In this manner, i descriptions are required to decode and view layer i . This can be generalised to:

$$\frac{L_i}{i}. \quad (3)$$

Once the sections are created, FEC is utilised to extend the layer data over the M descriptions, such that the higher the priority of the layer, the greater the level of error resilience, as illustrated in Figure 1b. In this manner, MDC-FEC contains an adaptive mechanism for description creation and error resilience allocation, but these mechanisms increase transmission cost proportionally to the level of FEC and is proportionally high compared to the initial level of SVC data, thus leading to a large increase in transmission cost relative to SVC. As can be seen in Figure 1b, the original SVC data are shown in blue while the FEC is shown in green. This image illustrates the increase in transmission cost with respect to FEC, relative to the specific SVC layer, but this image does not show the distribution of the respective SVC and FEC data per layer. Typically, FEC can provide either systematic or non-systematic encodings. Systematic schemes encode the original symbols as part of the transmitted stream, while non-systematic schemes encode and transmit the original symbols as new symbols. Raptor codes [47,48] propose that a systematic encoding, with encoded symbols interspersed among the original symbols, provides a greater level of decodability. Thus, in reality, each section contains a mixture of SVC and FEC data assuming a systematic scheme, while all data are FEC data in a non-systematic scheme. It is important to keep this thought in mind while reviewing description-based schemes.

To extend the section allocation over a number of frames per GOP, we write that an MDC description section from layer l from frame x , $S_{l,x}$, contains $\frac{L_{l,x}}{i}$ of the layer size, while a single *complete*

MDC description from frame x , as shown in Equation (4), contains the transmission cost of one section from layers 1 to N :

$$\sum_{l=1}^N \frac{L_{l,x}}{l}. \quad (4)$$

We view the total transmission cost of one complete MDC description, MDC_{D_c} , from each frame per GOP as

$$MDC_{D_c} = \sum_{frame=1}^{GOP} \left(\sum_{l=1}^N \frac{L_{l,frame}}{l} \right). \quad (5)$$

Note that it is not correct to multiply a single description by the number of frames per GOP, as each frame, as well as each layer, per GOP may have differing transmission cost, hence the requirement of the summation over all frames, using the $frame$ value, and the need to determine the layer cost per description section for each frame. Also note that the number of layers per frame, and number of frame rates per GOP depends on the underlying SVC encoding. In our equations for MDC, we determine the total transmission cost based on all layers required at the maximum frame rate. If a reduction in the frame rate is necessary, then a modified version of Equation (5) would mandate an additional variable, $frameStep$, which would increment over the frames not required. The following example illustrates a $frameStep$ of 2 that would half the frame rate. Note that the $frameStep$ value is dependent on the governing GOP value, such that the $frameStep$ value can never be larger than the GOP value and that the $frameStep$ value must always be a power of 2:

$$MDC_{D_c} = \sum_{frame=0, frameStep=2}^{GOP-1} \left(\sum_{l=1}^N \frac{L_{l,frame=frameStep*frame+1}}{l} \right). \quad (6)$$

Thus, the total transmission byte cost of MDC per GOP and at the maximum frame rate required to decode quality layer q can be seen as

$$MDC_{D(q)} = MDC_{D_c} * q. \quad (7)$$

The total FEC transmission cost overhead for MDC quality layer q can be characterised as $MDC_{D(q)}$ from Equation (7) minus $SVC(q)$ from Equation (1)

$$overhead = MDC_{D(q)} - SVC(q) \quad (8)$$

or using a simple single frame per GOP example

$$MDC_{D(q)} - \sum_{l=1}^q L_l. \quad (9)$$

Note that layer l defines a specific layer within the encoding and transmission of SVC, while quality, or layer quality, q defines the viewable quality achievable by decoding a number of descriptions.

Thus, to recap, MDC provides:

1. Increased transmission cost relative to SVC.
2. Increased error resiliency, proportional to the priority of the layer.
3. Viewable quality in some part reflective of network loss rate. Viewable quality is dependent on the number of descriptions available for decoding. Thus, as loss rates increase, the probability of a few descriptions being available for decoding increases.

4. Non-prioritised (equally important) descriptions. Each descriptions contains the same amount, i.e., number of sections, from each SVC layer.
5. Single description provides base layer decoding. Each description contains the base layer, thus minimum quality decoding is available once any description is received at the device.
6. High transmission cost for lower layer decoding. Each description contains one section from each layer, and, if a low quality layer is preferred, then receipt of sections from higher layers are not required but are received.

In the literature, several description-based streaming models have been proposed to reduce the transmission byte-cost of MDC or increase the achievable quality. These include:

- adjusting the levels of FEC, such as Adaptive FEC [49] and Enhanced Adaptive FEC [50],
- modifying the layer allocation per MDC description, such as transmitting the base layer as a separate MDC description [51]. The authors of [51] base their work on Priority Encoding Transmission (PET) [52], a prioritised packetisation scheme,
- modifying the base layer to create two individual descriptions [53],
- encoding one or more layers of an SVC stream into various bit rates, thus generating numerous descriptions composed of differing quality streams, such as Scalable Multiple Description Coding (SMDC) [54,55], and
- increasing the number of descriptions while reducing the byte allocation per description section of the SVC layer data and FEC, coupled with application-layer packetisation, such as Adaptive Layer Distribution (ALD) [56].

We focus on one of the description-based streaming models, namely ALD as it offers a comprehensive approach to description-based streaming and provides an ideal vehicle to illustrate the benefits of Streaming Classes for description-based schemes.

2.4. Adaptive Layer Distribution

Adaptive Layer Distribution (ALD) [56] has been proven to reduce transmission cost relative to MDC and provide consistent high levels of stream quality. ALD contains a number of components by which stream quality and transmission cost can be optimised for both network and user requirements. These include:

1. *Section Thinning Factor (STF)*: This component is utilised by ALD to increase the number of descriptions being transmitted, while decreasing the layer byte allocation per description *section*, thus decreasing overall stream transmission byte-cost. Figure 2 illustrates how the four MDC descriptions from Figure 1b, are increased to six ALD descriptions, using an STF value of 2. To illustrate how the transmission cost is reduced, from Figure 1b, four base layers were transmitted, consisting of one SVC data section (blue) and three FEC data sections (red), while, in Figure 2, when an STF of two is applied, and only two base layers are transmitted, i.e., the original SVC base layer is distributed over two additional ALD descriptions; thus, each base layer section is approximately 33.3% of the original SVC data. Apply the same percentage rate to the remaining base layer FEC sections and the total transmission cost of the base layer is 200% of the original SVC data, or two base layers (thus demonstrating a 50% reduction in transmission cost over MDC). Each layer of the original blue (dark) SVC data is partitioned over an additional two, STF value, descriptions, D_c-1 and D_c-2 , while the original level of error resilience (number of green (light) FEC sections per layer) is maintained. The additional ALD descriptions are shown in red, in order to illustrate a visual comparison in number of descriptions required by ALD, six, and MDC, four. The transmission byte-cost of devices requesting lower layer decoding is dependent on stream encoding and the ALD selection value for STF but is predominately larger than the relative transmission byte-cost of MDC.

2. *Section Distribution (SD)*: This component extends the benefit provided by equally important descriptions to the datagram level per frame. SD uses the well known Interleaving [57] technique to allocate a segment of each section per description to a datagram, as illustrated in Figure 3, as well as taking inspiration from the well-known Priority Encoding Transmission (PET) [52] prioritised packetisation scheme, thus limiting datagram loss to a segment of each ALD section, rather than being SVC layer or MDC description specific. This creates a coping mechanism for both single datagram loss and burst loss models. Thus, achievable quality per frame is based on the cumulative number of datagrams received, with stream consistency based on the level of datagram loss during transmission. One additional benefit of SD is that all datagrams per frame are of equal byte size.
3. *Improved Error Resiliency (IER)*: This component increases the level of FEC in the highest layers. As can be seen in Figures 1b and 2, layer 4 contains no FEC, such that the loss of any description results in an immediate downgrading of the stream quality. IER reduces the number of non-redundant sections of higher layers by distributing the higher layer data over one less section allowing for an additional FEC section. This minor increase in transmission byte-cost can yield a smooth play-out with fewer quality transitions at the highest layers.

L4.1	L4.2	L4.3	L4.4	L4.5	L4.6
L3.1	L3.2	L3.3	L3.4	L3.5	L3.6
L2.1	L2.2	L2.3	L2.4	L2.5	L2.6
BL.1	BL.2	BL.3	BL.4	BL.5	BL.6
Dc - 1	Dc - 2	Dc - 3	Dc - 4	Dc - 5	Dc - 6

Figure 2. An example of a four-layered SVC stream encoded as ALD, with an STF value of 2 (red denotes additional ALD descriptions, which contain existing SVC data).

Dc - 4	BL.4	L2.4	L3.4	L4.4
--------	------	------	------	------

Figure 3. ALD packetisation of D_c-4 from ALD in Figure 2. It can be seen that each ALD datagram contains section segments from all layers (red denotes packet header).

2.5. Synopsis

As we have seen, each of the scalable schemes contain known design issues. While the low transmission overhead is a benefit of SVC, the prioritised hierarchy and its dependency on the base layer are its greatest weaknesses. As we have highlighted, network transmission issues can affect all packets, and lower layer loss in SVC is detrimental to stream quality. MDC mitigates the effects of loss by overly increasing transmission cost, especially for users interested in lower layer video quality. More importantly, this transmission overhead represents a huge burden on users with limited bandwidth or device capabilities. This overhead is even more overwhelming when videos are encoded with a large number of layers to accommodate the existing diversity in mobile device capabilities. ALD provides the framework to achieve the high levels of adaptable stream quality promised by SVC, but the transmission byte-cost of devices requesting lower layer decoding is dependent on stream encoding and the ALD selection value for STF. However, ALD mandates a high level of transmission cost for devices requiring lower layer streaming.

In the following section, we present Streaming Classes as a novel transmission framework that eliminates the high transmission cost for description-based streams over limited bandwidth lossy links.

3. Streaming Classes

To alleviate these issues, we now propose Streaming Classes (SC), a novel transmission framework by which streamed media, either layered or description-based, are selectively grouped together to provide a means of reducing lower layer transmission byte-cost, while maintaining high levels of viewable quality for all users, irrespective of the layer requirements and original streaming model. In this manner, SC selects a number of layers from the underlying media stream, which best suit user requirements, and groups these layers together into *classes*. Figure 4 illustrates a six-layer SVC example, which is grouped as three hierarchical SC classes, where C1 denotes class one and contains the lowest layers, C2 denotes class two and contains the mid-range layers and C3 denotes class three and contains the highest layers. Similar to SVC, the SC classes contain a prioritised hierarchy, such that classes containing lower layers are needed to decode classes containing higher layers. For layered coding, SC will introduce a minor increase in the transmission cost to provide a sufficient level of resilience for all layers, thus adequately negating the effects of packet loss. However, for description-based models, SC will provide a marked reduction in transmission cost by selectively reducing the error resilience of the specific layers that are requested. For ALD, this will provide a means of reducing the lower layer transmission cost. In description-based encoding, the main transmission unit is the description, while, for layered models, the transmission unit is a layer. SCs reduce the transmission overhead by redefining the transmission unit as a selected set of section segments from different layers. As we have seen, selective packetisation of streaming data provides control over the effects of network loss on achievable quality, especially for the higher layers. With SC, we introduce this control to lower layers in the stream especially for devices, as previously mentioned, with limited bandwidth or device capabilities.

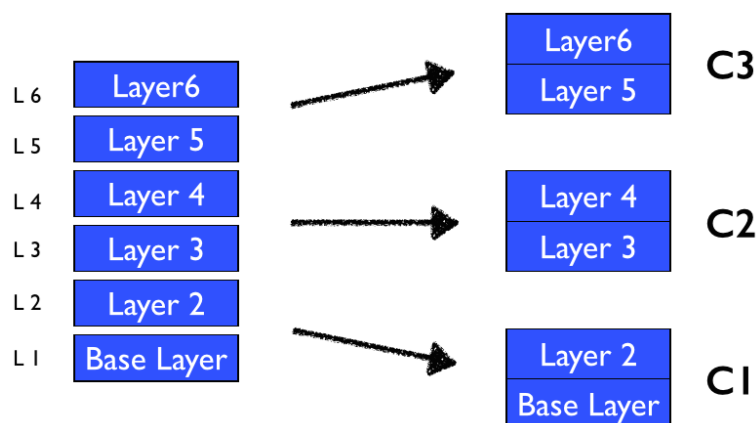


Figure 4. Example of a six-layer SVC stream grouped as three hierarchical classes.

SC is a hybrid mechanism, which uses the description structure of description-based models, and the prioritised hierarchy and layer intra-dependency of layered models to reduce transmission cost and maintain the viewable consistency of stream quality. SC is utilised to re-distribute the grouping mechanism of description-based models, i.e., one section from every layer in a description, and provides a mechanism to group layers and sufficient levels of error correction, in order to provide acceptable levels of transmission cost for all layers, irrespective of scalability. Thus, SC provides a means of grouping layers together to mitigate network loss and to reduce transmission cost for devices requesting a lower level of achievable quality. In this manner, and, by utilising the prioritised hierarchy of layered coding, SC provides classes based on the underlying combination of temporal, spatial and fidelity scalability, as well as the layer intra-dimensionality that is removed from description-based adaptive stream models.

4. Design Principles

The number of design options to consider can greatly increase the complexity, and the underlying decisions, mandated during encoding, packetisation, transmission, decoding and subsequent viewing (we call the process from encoding to subsequent viewing *stream flow*, while each of the individual steps from encoding to subsequent viewing we call *stream elements*). Examples of these design options include encoding requirements (number of layers, QP values, GOP size, etc.), expected network loss rates, video clip types, user quality selection and specific design options for the underlying streaming model (STF for ALD as an example), as well as the interaction and interdependency between all these design options. It is important to understand that decisions in one stream element will impact other stream elements, e.g., encoding decisions determine the stream bitrate, which changes the number of packets being transmitted, which in turn affects transmission issues, such as congestion and the loss rate, which influences the decodability of the stream and subsequently the level of viewable quality of the stream at the user's device.

Based on what we have learned from the evaluation of our initial stream models, SDC [58] and ALD [56], as well as the feedback from our subjective testing of scalable models [59], we view SC as more than just a means of reducing transmission cost for lower layer streaming, but as an architecture to allow one to appreciate that the achievable quality of a stream is dependent on so many factors. One design based on a number of defined variables can not solve all stream flow issues that can occur. Thus, the goal for SC is to consider all aspects of the stream flow and provide adaptive mechanisms for determining and supporting consistent high levels of viewable quality for all users. Table 2 presents the notation used in this section.

Table 2. Notation and definitions.

n	The number of SVC layers, L , per Group of Picture (GOP)
m	The number of SC classes per GOP
l	Integer value corresponding to the layer number of L_l
GOP_{number}	The number of frames per GOP
GOP_i	The i th GOP
μ	Mean or average network loss rate
μ_{GOP_i}	Loss rate for the i th GOP
μ_X	Loss rate for the duration of the stream. Also known as the Loss rate from GOP_1 to GOP_N , where N denotes the maximum number of frames per GOP
LR_{C_i}	Loss rate for Class i
PL_{C_i}	Packets lost for Class i
T_{C_i}	Packets transmitted for Class i
$\mu_{GOP_i}^{max}$	Determined maximum loss rate for the i th GOP

We begin by providing a brief overview of the five primary design principles for streaming classes:

1. Based on the observed levels of network loss, the *Class Packet Loss Rate* principle is used to determine an expected loss rate value, LR_{C_i} , for each class, C_i . The LR_{C_i} is used by each of the following four design principles. The goal of this principle is to mandate consistency of quality over time, e.g., reduce the frequency of quality change that occurs between adjacent frames, or GOPs, due to variations in network loss rate.
2. The *Layer Allocation and Hierarchy* principle defines how to group individual SVC layers together based on the interdependence that exists between different classes because of the inherent SVC layer prioritisation. The composition, or structure, of the individual classes is dependent on the initial SVC encoding, the GOP value selected, user requirements, and the defined combination of spatial, temporal and quality scalability. One example of layer allocation is to group lower layers together into a single class, so as to reduce transmission cost for devices requesting lower quality streaming or for devices on constrained networks.

3. The principle of *Error Resiliency* is utilised by each class to add sufficient levels of error resilience to every layer to combat defined levels of network loss. To reduce overall transmission cost, we implement an inverse relationship between layer level and error resilience, i.e., the higher the layer level, the lower the error resilience. This relationship provides the more important layers per class with higher levels of resiliency.
4. Similar to layered and description-based models, SCs are created with a hierarchical nature such that the more SCs received by the user, the better the quality of the decoded video. The principle of *Streaming Class Structure* defines each group of layers as a prioritised class, i.e., classes containing lower layers are more important to viewable quality than classes containing higher layers, as higher layers in the SVC hierarchy are predisposed to be dependent on lower layers. Thus, the level of viewable quality is dependent on the number of classes received, as well as the number of decodable layers per class.
5. The principle of *Class Packetisation and Granularity of Packet Data Byte-size* is utilised so as to provide layer equality per class, i.e., by utilising Section Distribution (SD), each packet per class contains a segment of every layer per class, consequently viewable quality is dependent on the number of packets lost rather than on the contents of the packets that are lost.

Thus, these design principles provide a selection of options by which both layered and description-based models can increase viewable quality. These options range from:

1. Encoding decisions such as GOP value, and spatial, temporal and quality scalability.
2. Transmission choices such as selective packetisation, the allocation of layers to classes and intuitive adaptation techniques, which adjust to network conditions e.g., packet loss and error resiliency allocation.
3. User requirements such as the quality of the media clip requested, and the corresponding maximum layer value and associated lower layer dependencies required to decode the requested quality.

In the following sections, each of the five primary design principles shall be examined in further detail.

4.1. Class Packet Loss Rate

The principle of “Class Packet Loss Rate” is utilised to determine the loss rate of each class, LR_{C_i} , based on the overall loss rate in the network. The packet loss rate for media streaming is primarily based on the number of packets unavailable during decoding of a specific frame or GOP. The unavailability of a packet during decoding can be caused by numerous issues; these include packet delay [60] (due to data transmission over multiple network links), buffering [61] (bottlenecks in the network routers) and congestion; both rate control [62] (slowing throughput to reduce loss) and packet dropping [63] (non transmission of packets to reduce loss). In general, the percentage of packets lost in the network will not equate to the same percentage of lost transmitted data, as the transmitted data tends to be composed of different sized packets and the percentage of lost transmitted data is dependent on the distinct packets that are lost. It is important to consider this relationship when we examine the effects of network loss.

In the literature, numerous mechanisms exist to determine the current network loss rate, example of these include Server-side such as Realnetwork Helix Mobile Media Server Rate Control [64], Web Server TCP packet loss determination for Quality of Service (QoS) [65] and IBM Unix network Performance analysis [66], Client-side such as Quality-Oriented Adaptation Scheme (QOAS) [67], Application-level Estimation [68], and NetPolice: loss rates in ISPs [69], from nodes within the network, such as iPlane: an information plane for distributed services [70], RON: resilient overlay networks [71], and Queen: estimating packet loss between arbitrary internet hosts [72], as well as from SDN control [73,74].

For our adaptation to network loss, we need to consider two issues:

- i. How the variation in network loss over the duration of a media clip influences the loss rate for any given individual GOP:

If we were to assume that the network loss rate is consistent over the duration of a media clip, then, by simply transmitting a percentage of redundant media proportional to the average network loss rate, this would enable consistency of viewable quality. However, in reality, variations in network loss rate will occur. Each GOP provides adequate statistical data to infer the current loss rate per GOP as well as the overall variation of viewable quality. Thus, for each GOP, GOP_i , we first determine the expected network loss rate, μ . μ is the mean or average loss rate over all previously transmitted GOP. This mean value does not mandate that all future GOP would encounter the same μ value but infers that future GOP shall include redundancy that accommodates to μ . The value of μ is affected by both the loss of an individual packet and the loss of groups of adjoining packets, known as bursty loss (we can define all loss as bursty loss if we include the loss of a single packet as a burst rate of 1) and as such the loss rate of the next GOP can range from 0 to 100% depending on the level of bursty loss and the number of frames per GOP.

In a binomial distribution, the variation that occurs from the mean value is limited in range. In network loss, the instantaneous GOP loss rate can range from 0 to 100% but when averaged out and taking minimum and maximum values, the variation is also limited in scope (between adjacent GOP). It is important to note that some outlying minimum and maximum values can occur, but these values are dependent on a number of factors, such as number of frames per GOP and the levels of traffic over the transmission network. In our evaluation, we illustrate how the variation in packet loss that can occur per GOP is dependent on the number of frames per GOP. Over the duration of the media clip, we determine the minimum and maximum packet loss value per GOP and show how the interval between these packet loss values decreases as the number of frames per GOP increases.

If we assume that the variation that can occur is limited in scope and is similar in nature to a binomial distribution, we can utilise σ_μ , which equates to the square root of the variance to cap the level of error correction required. In this manner, μ and σ_μ are used to define a statistical maximum threshold on the expected network loss rate for the current GOP, which we define as the maximum expected loss rate, μ^{\max} , for a given GOP_i or $\mu_{GOP_i}^{\max}$. $\mu_{GOP_i}^{\max}$ limits the increase in transmission cost while providing an acceptable level of adaptive error resiliency.

- ii. How the loss rate for an individual GOP determines the loss rate per class:

The viewable quality of a single non scalable stream, e.g., one resolution, one frame rate, and one quality level, is based on the average loss rate μ over a predetermined length of time, e.g., from the start of the stream to the current point in time, while, for scalable streams, e.g., layered/description-based models, the viewable quality is dependent on the average loss rate, μ , over all transmission units (layer, description, section, packet-dependent on the underlying streaming model), e.g., the distribution of μ over each of the transmission units and how this allocation of loss affects the decoding of the underlying layered data. Thus, the defined loss rate for each individual transmission unit is based on the percentage loss rates of all other transmission units and μ . This is also true for each class in SC. The μ is mean over all classes and does not mandate the loss rate of each individual class will equal μ .

Hence, the overall loss rate, with respect to each class, can be defined as the summation over the number of packets lost per class divided by the number of packets transmitted per class divided by the number of classes or

$$\mu = \frac{\sum_{i=1}^m \frac{PL_{C_i}}{T_{C_i}}}{m}. \quad (10)$$

It can be seen that, for a given loss rate over a number of classes, the average loss rate per class can vary from zero loss, to the transmission cost of all classes multiplied by the overall network loss rate, e.g., assuming three classes of equal transmission cost, a μ of 10% could mandate an individual class loss rate, LR_{C_i} , of between 0, minimum loss rate $LR_{C_i}^{\min}$, and 30%, maximum loss rate $LR_{C_i}^{\max}$. If we again assumed three classes and 10% loss, where C_1 has a transmission cost of 100 bytes, C_2 has a 1000 byte cost and C_3 has a 10,000 byte cost. Then, the loss rates range from a minimum of 0 for all classes to a maximum of 11.1% for C_3 and 100% loss for both C_1 and C_2 . Total transmission loss at 10% is 1110 bytes and this is larger than C_1 and C_2 combined. The effects of individual class loss rates are investigated in the *error resiliency* principle and examples of the variation that can occur between classes of the same GOP are illustrated in the evaluation section of this section.

For each GOP, GOP_i , the loss rate of each class, LR_{C_i} , is passed as an argument to the other four primary design principles. μ can be determined from these loss rates or can be passed as a separate argument.

- i. For layer allocation, this value will assist in determining the layer distribution to classes during periods of high loss rate or bursty loss.
- ii. For error resiliency, this value will determine the forward error correction (FEC) allocation rate for the highest layer in each class.
- iii. For class structure, this value will assist in determining the number of classes being transmitted. During moments of high network loss, classes containing higher byte-cost layers may be dropped, so as to reduce congestion.
- iv. For class packetisation, this value will assist in determining the byte-size and layer segment allocation per packet.

Each of these items will be further explained in their corresponding section.

4.2. Layer Allocation and Hierarchy

The definition of streaming classes would vary depending on the scalability techniques adopted during video encoding. To illustrate the design options of defining SCs, we consider a generally encoded video using a combination of spatial, temporal, and quality scalability for the well-known crew video [75].

The output of the encoder is shown in Table 3 in which the (D,T,Q) tuple, (D for dependency_id-spatial resolution dependency), T for temporal_level and Q for quality_level) represents the level of dependency in the three scalability dimensions. A higher index in any of these fields indicates that decoding the corresponding layer requires receiving all previous layers with a lower index in the same field. For example, decoding layer 7 (1,1,0) implies that we should receive all layers including lower values in the D and T fields; i.e., the streaming client should also receive layers 6 (1,0,0), 1 (0,1,0), and layer 0 (0,0,0), assuming that inter-layer dependency is enabled in the encoding process (CGS selection). It is worth noting that, if scalability is performed in one dimension (or Medium Grain Scalability (MGS) encoding selection), the reception of a higher quality video mandates the reception of all lower layers. For example, layers 0, 1, and 2 may be considered as temporally scaled video with Quarter Common Intermediate Format (QCIF) resolution (176×144)

and one quality level. Hence, for videos encoded with a single scalability dimension, there exists one option for layer aggregation, which is grouping subsequent lower layers.

Table 3. Encoder output for a general encoding scheme for the crew video, composed of two resolutions, two quality levels and three frame rates.

Layer	Resolution	Frame Rate	(D,T,Q)
0	176 × 144	7.5	(0,0,0)
1	176 × 144	15	(0,1,0)
2	176 × 144	30	(0,2,0)
3	176 × 144	7.5	(0,0,1)
4	176 × 144	15	(0,1,1)
5	176 × 144	30	(0,2,1)
6	352 × 288	7.5	(1,0,0)
7	352 × 288	15	(1,1,0)
8	352 × 288	30	(1,2,0)
9	352 × 288	7.5	(1,0,1)
10	352 × 288	15	(1,1,1)

The presence of two types of scalability in the encoded video creates more design options as shown in the previously presented combined quality-resolution scalability. If these two dimensions include a temporal dimension, the layer grouping can be performed in different ways. To illustrate, consider layers 0 to 5 in Table 3. This subset of layers can be considered an example for a combined temporal-quality encoded video at QCIF resolution. In this case, two basic design options are envisioned for grouping:

- SCs are created from layers having the same quality index. For example, layers 0 to 2 and layers 3 to 5 represents two streaming classes respectively.
- SCs are created based on frame rate (temporal dimension). For example, grouping layers having 7.5 fps to form the lowest SC, layers played at 15 fps as a second streaming class and layers played at 30 fps as a third streaming class.

Which option to choose would depend on the underlying bitrate of the video and the individual user's requirements. The transmission cost of both options is the same, as an equal number of layers are transmitted by both options. Option one can only vary the frame rate once all classes are received as layer allocation is based on quality levels, while the second option is more adaptive to users requesting lower frame rates, as each frame rate is contained within a distinct class.

Similarly, a video encoded with temporal and resolution dimensions would have two basic design options for creating the SCs. To illustrate, let layers 0 to 2 and layers 6 to 8 represent a video encoded with temporal and resolution scalability. Hence, the two design options would be aggregating based on frame rate or based on video size. The former would also be favourable because the latter would produce a high data rate in the lower streaming classes. Additionally, the former would be more appealing as the higher frame rate in the aggregated lowest class would result in a better quality video. Last but not least, consider an encoded video over the three scalability dimensions as shown in Table 3 and note the many SC aggregation options that would be possible. However, as established in our previous presentation, aggregating layers with different resolutions contradicts with the principals of SC (increasing transmission cost for lower layer streaming). Hence, aggregation over the temporal and quality scales remain the two possible options.

To this end, the layer aggregation criteria may include more than one rule. For example, the aggregation of several layers over one of the scalability dimensions can be accompanied with another constraint on other factors such as the resultant transmission byte-cost of the aggregated layers. Another possibility is that the aggregation criterion considers target bit-rates matching existing

bit-rates in real networks. To further illustrate, if the aggregation of the three quality layers for a single resolution would result in a data rate exceeding a target bit-rate, one can alternatively use more than one streaming class for this set of layers by allocating the lowest of them as one streaming class and the remaining two layers as a second streaming class. The proposed one-two splitting is suggested to create a balanced set of streaming classes.

As illustrated in the adaptation to network loss section, we can also consider the underlying loss rate in the network while determining the layer allocation per class. The effects of congestion can be reduced, in some part, by a reduction in stream bit-rate. Thus, by removing the allocation of high byte-cost layers, we can reduce the level of network loss. In addition, careful consideration is needed when we consider the prioritised hierarchy of the classes, e.g., how important is the content of this class, and as such the maximum achievable quality of this class, to all other classes.

Thus, to recap, the number of layers and layer allocation per class can be based on a number of options, these include but are not limited to:

- i. *adjacent layers*: an i number of adjacent layers. Subsetting the total number of SVC layers into classes, where each class contains an equal number of layers.
- ii. *temporal, spatial or fidelity layering*: creating classes based on distinct frame rates, resolutions, quality levels, or combinations of same. Combining higher layer data with lower layer data may be counter productive to the reduced transmission cost objective of SC.
- iii. *layer intra-dependency*: the intra-layer dependency of requested layers.
- iv. *stream layer selection per requesting domain/network location*: the layers being requested per individual domain or co-located networks. Classes can be created for specific domains, or network regions, which reduce overall transmission cost to the specific layers requested by these geographical locations.
- v. *the number of distinct quality layers being requested*: the number of distinct layers being requested. Classes can be created which are based on the specific layers required to decode quality layer q . Consequently, removing the cost of transmitting un-required layers.
- vi. *the number of users per quality layer*: grouping the most requested layers together. Creating classes based on overall user requirements. The greater the requests for a given layer, the higher the class prioritisation, or lower class number, for that layer.
- vii. *based on bit-rate selection*: thresholds of available bandwidth and cumulative transmission byte-cost. Creating classes based on network throughput, constrained links or overall congestion rates. While not referencing scalable video specifically, but commenting on the reliable transmission of TCP traffic and using layered video to illustrate adaptation of quality dependent on available bandwidth, the TCP Rate Adaptation Protocol (RAP) proposed in [76] is a seminal paper in this regard. The authors of [76] investigate the rate-adaptation of layered media dependent on the timescale of round-trip times but maintain viewable quality for longer periods of time by using buffering to accommodate mismatches between transmission and consumption rates.
- viii. *replication of layer data*: it is not infeasible that numerous classes may benefit from the allocation of the same layer, e.g., allocation of the base layer to numerous classes, with each of the classes based on a different spatial, temporal or quality scalability. It may be more cost effective, with respect to transmission cost for individual users, to transmit the individual layer as a single class, or allocated to numerous classes, so as to reduce the receipt of higher layers from unwanted dimensionality.

As previously stated, the goal of SC is to consider all aspects of the stream flow and then choose which option(s) maximise overall viewable quality. Thus, which “Layer Allocation and Hierarchy” option(s) to choose at runtime will be governed by the effects of the chosen option(s) on the other design principles and overall stream quality.

4.3. Error Resiliency

Irrespective of transmission unit, the transmission cost overhead between layered and description-based models is fundamentally the level of error resilience allocated. Thus, the goal of the error resiliency principle is to provide balance between transmission cost and achievable quality. As previously described, section distribution (SD) offers a means of delivering packet equality for description-based models, such as MDC and ALD. It is this packet equality and the incremental levels of error allocated per SVC layer, which provides consistency of viewable quality and mandates that viewable quality is dependent on the number of packets lost rather than on the contents of the packets that are lost. By implementing SD in the streaming class model, we can mandate that the number of decodable layers for class i , C_i is dependent on the level of error allocation per layer and the corresponding class loss rate, LR_{C_i} .

We have already seen how the loss rate of class i , LR_{C_i} is dependent on both the μ and the loss rates of all other transmitted classes. Hence, the level of error resiliency per layer must take into consideration:

- i. the priority of the layer in the original SVC hierarchy. Lower layers per class contain higher level of error correction. This allocation of error correction provides an incremental increase in quality, dependent on the level of network loss, which is consistent with description-based streaming models and also a graceful degradation in viewable quality as network loss increases.
- ii. the variation that can occur in both μ and the loss rates of all other transmitted classes. As seen, the variation that can occur between classes for each GOP is relative to the loss rate of all classes, thus this variance, or deviation, must be taken into consideration.
- iii. the maximised consistency of quality for this class and for all classes that are dependent on this class. In SVC, higher layers are dependent on the achievable quality of lower layers and SC is similar in this regard. Dependent on the allocation structure of the classes, higher classes will predominately be accessible only when all layers in a lower dependent class are decodable. To benefit from this dependency, all layers per class must contain error resiliency, and higher levels of error resiliency may be required in lower classes to maximise quality for all classes.
- iv. the bitrate of the allocated layers. Low bit rates per layer, which can occur with larger GOP values and for lower frame rates, may require more than the minimum error resiliency allocation to recover from loss, as the byte allocation per packet may be too small.
- v. the addition of increased error resiliency will impact on transmission cost, viewable quality and network loss. As we determine error resiliency based on predetermined loss rates, will increasing the error resiliency per layer impact the current levels of network loss? We can consider three examples that illustrate the varying levels of error allocation, relative to current network loss rate:

- i. Define the error allocation rate for the highest layer in class i to be equal to the current level of network loss, i.e., LR_{C_i} .

This option will, in some instances, only covers lower loss rates than the current rate, e.g., assuming a loss rate of 10%, if we add 10% additional data to the layer contents, we effectively increase the transmission cost to 110%. Assuming the loss rate does not increase with increased traffic, then the effects of the same 10% loss rate will equate to 11% of the 110% being transmitted, consequently only 99% is decodable and the highest layer is un-viewable. Note that this disparity will be dependent on the packetisation rate, as we shall see later. If the increased error rate can be allocated to the same number of packets as the original layer data, then even with 10% of the packets lost, the layer is still decodable.

- ii. Define the error allocation rate for the highest layer to be equal to the current level of network loss plus the standard deviation that can occur for this loss rate, i.e., $\lceil LR_{C_i} + \sqrt{LR_{C_i}} \rceil$. We define this as the allocated loss rate value for the maximum viewable

quality or LR_{\max} . This would equate to 10% + 3.17%, respectively, or a rounded up value of 14% FEC for the highest layer, assuming a 10% packet loss rate.

This higher level of error allocation covers some of the packet loss rate variation over time but may still be susceptible to loss during high bursty conditions.

- iii. Define the error allocation rate for the highest layer to be upper bounded by a maximum threshold based on the expected network loss rate required to recover from the current level of network loss, which we will define as $LR_{C_i}^{\max}$. This loss rate will need to consider both the initial loss rate, LR_{C_i} , and the loss rate that will affect our newly added FEC, $\sqrt{LR_{C_i}}$. An algorithm to determine this value is provided in Equation (11). $LR_{C_i}^{\max}$ denotes the current determined loss rate for the maximum viewable quality $\lceil LR_{C_i} + \sqrt{LR_{C_i}} \rceil$:

$$LR_{C_i}^{\max} \geq \frac{LR_{\max}}{1 - LR_{\max}}. \quad (11)$$

Finally, as per our example, if we assume LR_{\max} is 14%, we can determine the minimum additional FEC value required to recover for this level of maximum packet loss, as shown in Equation (12):

$$\begin{aligned} LR_{C_i}^{\max} &\geq \frac{14\%}{1 - 14\%}, \\ LR_{C_i}^{\max} &\geq \frac{14\%}{86\%}, \\ 17\% &\geq \frac{14\%}{86\%}. \end{aligned} \quad (12)$$

Such that, for a packet loss rate of 10%, a standard deviation of 4% and a minimum recovery rate (based on packet loss + standard deviation) of 3%, we can define a $LR_{C_i}^{\max}$ of 17% FEC.

This option covers both the variation in packetisation rate as well as minor variations in network loss rate.

- vi. the original streaming model. For SVC, the transmission unit is a complete layer and the packetisation options outlined above take into consideration the entire layer plus allocated error resiliency. Thus, small incremental increases in error resiliency are easily accommodated into the packetisation option. For description-based models, the transportation unit is either a description (MDC), a section (SDP) or a packet (ALD), such that the packetisation is defined based on the size of the underlying unit. Additional error resiliency is then based on the addition of a predetermined section size based on the value of a layer index (plus STF for ALD). These section sizes are static in nature and the addition of a section adds a pre-defined level of increased transmission cost to a class, e.g., for ALD, assuming five layers and an STF of 3, this defines that each section in the highest layer, layer 8, contains 12.5% of the overall transmission cost of that layer. As ALD adds additional sections to define the structure of increased error resiliency, each increase in section for layer 8 will mandate an increase of 12.5% of the layer byte size. Minor increases can be achieved by using Improved Error Resiliency (IER), but even then the increase in overall cost can be quite high and possibly higher than the required LR_{C_i} percentage. A minor increase in error resiliency can be allocated, but, as ALD uses sections sizes to define SD packetisation, this increase must be distributed over all sections or must equal the defined SD packetisation sizes.

To sum up: the allocation of error correction must take into consideration the structure of the original stream model. With layered streaming, the increase in error correction can directly equate to the expected network loss rate, while, for description-based models, the description structure is defined with default levels of error resiliency, e.g., based on predefined section sizes, which may overly increase error correction levels and thus elevate transmission cost.

Thus, to recap, for each class; each layer must contain a minimum level of error resiliency, consistent to the LR_{C_i} determined by the “Class Packet Loss Rate” principle. Lower layers per class must contain incrementally higher levels of resiliency, to counteract the variation that can occur due to loss rates in other classes, thus providing graceful degradation in viewable quality as loss increases. Examples of allocation rates for LR_{\max} and $LR_{C_i}^{\max}$ based on a packet loss rate of 10% are presented in Appendix A.

Once the error allocation rate for the highest layer per class is determined, the incremental increase in error allocation per lower layer must be determined. This can be defined on a fixed incremental increase, a variable level of increase (dependent on external factors, such as packetisation rate, layer allocation, network conditions, user requirements to name but a few), or can be based on the allocation rate of the adjacent higher layer plus a determined level of standard deviation, as used in our SVC evaluation results in Appendix A.

4.4. Streaming Class Structure

Similar to the layer structure in SVC, SC imposes a prioritised hierarchy based on the importance of the layers contained within each class. This class hierarchy defines which subset of classes is required to decode a requested layer. In this section, we explore how variations in class composition can be utilised to maximise viewable quality per class. Dependent on layer allocation, the viewable quality of higher classes are dependent on the higher prioritised lower classes. Such that the loss of a lower layer will impact the viewable quality of all or a subset, dependent on encoding dimensionally, of the higher layers for all classes. Thus, the goal of this principle is to maximise quality in all lower classes.

As we have seen, the loss rate per class can vary dependent on the overall loss rate, μ , and the loss rate of the other classes. While error resiliency provides a mechanism to allocated acceptable levels of error correction to individual layers, there exists a balance between sufficient error resiliency and increases in overall transmission cost.

We offer two class composition options by which to maximise viewable quality per class. Both options illustrate the balance between error resiliency and transmission cost and are based on the interdependence of the classes. For ease of illustration, we present an example based on ALD, and note that the same class composition option would also hold true for SVC and MDC (remember ALD with STF = 0 is MDC). Figure 5 illustrates a six-layer ALD stream with an STF of three. We shall create three classes, based on layer adjacency, such that class 1 (C1) is composed of the base layer and layer 2, class 2 (C2) is composed of layer 3 and layer 4, and class 3 (C3) is composed of layer 5 and layer 6. The two class composition options, based on the defined class hierarchy, can be defined as:

- i. *Independent Class Composition (ICC)*: With ICC, layer data of a class is contained only within the class. This is illustrated in Figure 6, where we can see that the layer data for each class is contained only within the defined class. In this instance, the error allocation rate, as determined by the “Error Resiliency” principle, governs the overall effects of variations in the network loss. A higher level of error resiliency in the lower classes may be sufficient to recover from loss. ICC maintains packet equality per class, mitigates loss by allocating sufficient levels of error resiliency and maintains acceptable levels of transmission cost per class. Examples of ICC based on a packet loss rate of 10% are presented in Section 4.5 and Appendix A.
- ii. *Increased Class Interdependency (ICI)*: ICI is an enhanced version of ICC, where additional lower class data is allocated to the higher classes to increase the viewable quality of the lower classes, and thus increase overall quality of the higher classes. The initial error allocation rate, as determined by the “Error Resiliency” principle, is utilised by ICC, and then ICI assigns additional unallocated layer data from lower classes to the higher classes. This is illustrated in Figure 7, where we can see that two sections from the Base Layer and layer 2 are allocated to class 2 (C2), while a single section from the four lower layers are allocated to class 3 (C3), thus allocating all of the remaining FEC sections from Figure 5 across the higher classes. Note how the number of FEC sections allocated to the higher classes is consistent with the number of FEC sections allocated to the lowest

layer of each higher class. While this example illustrates FEC allocation for description-based streaming, the same steps would be used for layered streaming, but the allocation rate would be based on the percentage of FEC in the lowest layer rather than a defined number of sections. Note how the FEC and layer data allocation in C1 in Figure 7 (ICI) and Figure 6 (ICC) are identical, as no additional data are allocated to the lowest class. ICC is beneficial when variation in the network leads to loss rates greater than the error allocation rate for a given lower class layer. In this manner, the availability of lower layer data in a higher class may provide sufficient additional data to maximise the given lower layer. In this option, in addition to the lower layer data allocated to the higher classes, a lower level of error resiliency allocated to the lower classes may be sufficient to recover from expected network loss. A minor increase in transmission cost, relative to the level of additional lower class data, is to be expected.

Finally, the level of lower layer data allocated to the higher classes permits the higher classes to experience an equivalent level of loss in the lower classes while still permitting complete decoding of the higher class. The level of lower layer FEC allocated to the higher classes is adaptive to the needs of the stream flow. Examples of this FEC level include adaptation based on the levels of loss in the network, and it may reflect the priority of a given lower layer or may be dependent on consistent decoding of a given higher layer.

Packet equality is maintained in the higher classes, as the lower layer data are distributed over all transmitted packets. For lower class subscribers, one additional benefit from the ICI composed higher classes is that, if the achievable quality in a lower class is marginally less than maximum, then the lower class user can subscribe to a subset of the higher class packets, thus improving quality by receiving segments of lower layer data, while marginally increasing transmission cost by receiving segments of un-required layers. As ICI mandates that all higher classes contain lower class data, the lower class user can select packets from all higher classes, therefore benefiting from real-time data acquisition of lower layer data, without the delay of retransmission, which is highly beneficial to media streaming.

If the error resiliency allocation rate of ICC is sufficient in all classes, then the ICI allocation of lower class data to higher classes is an unacceptable and un-required increase in transmission cost.

L6.1	L6.2	L6.3	L6.4	L6.5	L6.6	L6.7	L6.8	L6.9
L5.1	L5.2	L5.3	L5.4	L5.5	L5.6	L5.7	L5.8	L5.9
L4.1	L4.2	L4.3	L4.4	L4.5	L4.6	L4.7	L4.8	L4.9
L3.1	L3.2	L3.3	L3.4	L3.5	L3.6	L3.7	L3.8	L3.9
L2.1	L2.2	L2.3	L2.4	L2.5	L2.6	L2.7	L2.8	L2.9
BL.1	BL.2	BL.3	BL.4	BL.5	BL.6	BL.7	BL.8	BL.9
Dc 1	Dc 2	Dc 3	Dc 4	Dc 5	Dc 6	Dc 7	Dc 8	Dc 9

Figure 5. ALD with six-layer and an STF of 3.



Figure 6. ALD streaming classes using the ICC class composition option—C1 denotes class one, C2 denotes class two and C3 denotes class three.

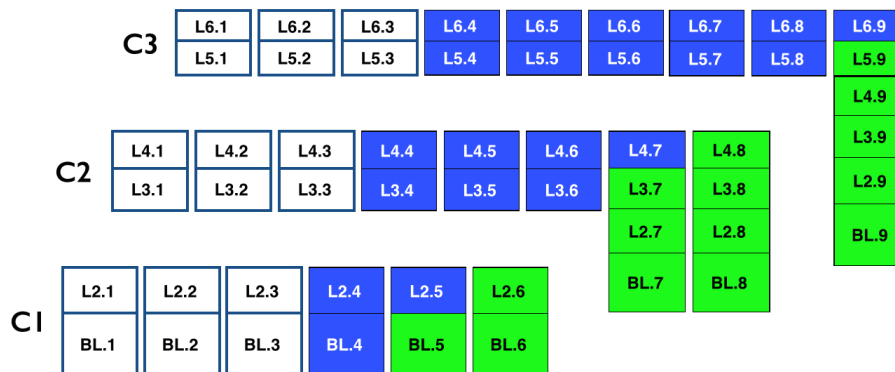


Figure 7. ALD streaming classes using the ICI class composition option—C1 denotes class one, C2 denotes class two and C3 denotes class three.

4.5. Class Packetisation and Granularity of Packet Data Byte-Size

As we have seen, Section Distribution (SD) was initially designed to extend the concept of equal importance from description to the packet level per frame. SD allocates a segment of each section per description to a packet, as a result limiting packet loss to a portion of each description section, rather than being SVC layer or MDC description specific. In this manner, SD creates a coping mechanism for both single packet loss and burst loss models. Thus, achievable quality per frame is based on the cumulative number of received packets, with achievable quality directly reflecting the level of packet loss during transmission.

One additional benefit of SD is that all packets per frame are of equal byte size. This equality is provided in both packet byte-size and packet priority. In addition, as the number of frames per GOP increases, SD will provide data equality for all frames within the GOP. In [21], the authors highlight that packets of dissimilar processing times, produce dissimilar transmission times. Such that, by maintaining such packet byte-size equality, the order of packet delivery is improved. Hence, SD packet equality improves consistent delivery in network transmission.

While SD provides packet equality based on an equal allocation of segments for each layer per class, to provide increased resiliency to network loss, we also consider the byte-size of each packet to determine if a byte-size threshold is beneficial to achievable quality. Let us assume for simplicity that a single frame contains eight SVC layers, and each layer is the same byte-size, e.g., 540 bytes. Total transmission cost for this frame is 4320 bytes, which would equate to a minimum of three packets, assuming a data size of 1440 bytes per packet and the layers are allocated in increasing order until the packet data-sizes are full. Let us also assume a network loss rate of 10%, which, in our example, would equate to one lost packet or over 2.5 lost layers. Best case (B-C) scenario is where the data from packet

three, e.g., layers 6, 7 and 8 were lost and layer 5 can be decoded, while worst case (W-C) scenario is where packet one, e.g., the base layer, layer 2 and 3 were lost and the frame is un-decodable. With the current packet data byte-size allocation, 1440 bytes, we incur the same level of layer loss for all packet loss rates from 1% to 33% inclusive.

By increasing the number of packets transmitting the layers, we can reduce the effects of the layer loss relative to the current loss rate. By increasing the number of packets to eight, we incur a minor increase in transmission cost, relative to the increased number of headers, while the effects of the 10% network loss is reduced to one packet, e.g., one specific layer. Each packet now contains 540 bytes, or one complete layer, and a single packet can incur a loss rate of between 1% to 12.5%. B-C for this simple decision increases viewable quality to layer 7, while in, W-C, the frame is still un-decodable. With eight packets, an optimal threshold for number of packets has been reached. An increase in the packet numbers will not increase viewable quality, as each packet will subsequently contain either segments of multiple layers or a segment of a single layer, and the loss of a segment will negate decoding of the entire layer.

Consequently, we present two packetisation options for SCs, namely reduced overhead packetisation and improved resiliency packetisation.

1. *Reduced Overhead Packetisation (ROP)* In this scheme, the data belonging to the same layer (including FEC sections) in each SC are aggregated to create one super section per layer. For SVC, this would combine layer and FEC data, while, for description-based models, individual description sections and associated FEC sections would be merged. These super sections are then packetised using the SD mechanism, where each packet contains a portion of the super section from each layer in the streaming class. ROP tends to create packets with large data byte-size content. However, the loss of any packet typically increases the probability of large amounts of multiple layer data being lost, thus reducing the availability of layer data by which stream quality is maximised. Example: C_1 shown in Figure 8a illustrates two SVC layers and associated FEC allocated to a single class. Let us assume that the total transmission cost is 7200 bytes. With ROP, we treat each layer per class as one super section and packetise using the minimum number of packets, e.g., five packets at 1440 bytes. Thus, the loss of a single packet equates to 1440 bytes of lost data from layers BL and L2. Should we have used a description-based model for this examples, as per Figure 8b, ROP would have combined the individual sections illustrated in Figure 8b, so as to represent Figure 8a and packetisation would be identical to the previous example.
2. *Improved Resiliency Packetisation (IRP)* For description-based streaming, we are able to use the section sizes of the underlying descriptions to define the governing threshold byte-size allocation of our packetisation. As per Figure 8b, SD is applied to the individual sections of both layers. Noting that the number of packets depends on individual section sizes in comparison to the super section sizes in ROP, more packets are typically generated on using this scheme.

For SVC plus error resiliency, we have no easily selectable segmentation markers. For this, we use the packetisation threshold as previously defined, by which we can specify our SD packetisation requirements. Figure 8c illustrates the packetisation of Figure 8a based on a threshold value, which determines the maximum combined byte-size of both layers allocated to a single packet.

As can be seen by using IRP, the loss of one packet would affect a smaller portion of the class data. Note that both ROP and IRP would benefit from SD but IRP has a better granularity. The main drawback of IRP is the additional overhead associated with every packet. Example: again let us assume that the total transmission cost of C_1 shown in Figure 8b is 7200 bytes. With IRP, and for description-based models, we maintain the descriptions structure, e.g., eight descriptions, and packetise based on the byte size of each description, e.g., 900 bytes. Thus, the loss of a single packet equates to 900 bytes of lost data from layers BL and L2, thus reducing the effects of packet

loss on achievable quality by increasing the number of packets transmitted. An example of IRP packetisation of SVC in Figure 8a is shown in Figure 8c. Note the data between the black vertical lines denote the packetisation segments of each layer.

Using the IRP packetisation example, we note that, typically, the byte-cost allocated per packet is less than the maximum byte-size of the packet data, circa 1440 bytes. Thus, one additional option to increase viewable quality is to allocate levels of error resiliency to the layers, such that the total layer cost is less than or equal to the maximum byte-size of the packet data. Therefore, we do not increase the number of packets transmitted but only the data content of the packets.

As an example, we assume the transmission of eight packets and an initial error resiliency allocation rate of 14% for the highest layer. Assuming a 2% increase in FEC for each lower layer, this provides an error resiliency allocation rate of 14% to 28%. We then take the eight layers and their respective error resiliency rates and packetise using SD. Now, each packet contains a segment of every layer, or approximately 12.5% of each layer plus a percentage of the initial layer error resiliency allocation. As the highest layer contains 14% error resiliency and each subsequent layer contains a higher level of error resiliency, both B-C and W-C with a network loss rate of 10% mandates layer 8, or maximum layer, decodability.

This is a simple example, but the underlying concept holds true for layers with differing byte-sizes, for larger GOP, for increased numbers of packets and for varying error resiliency allocations. Once the threshold for packet sizes/number, as well as error resiliency allocation rates, is defined, we can then begin to offset optimality against overall transmission cost to define a balance between viewable quality and the effects of network loss. As illustrated by the error resiliency allocation rates, we can always implement IER to further increase the resiliency of select layers.

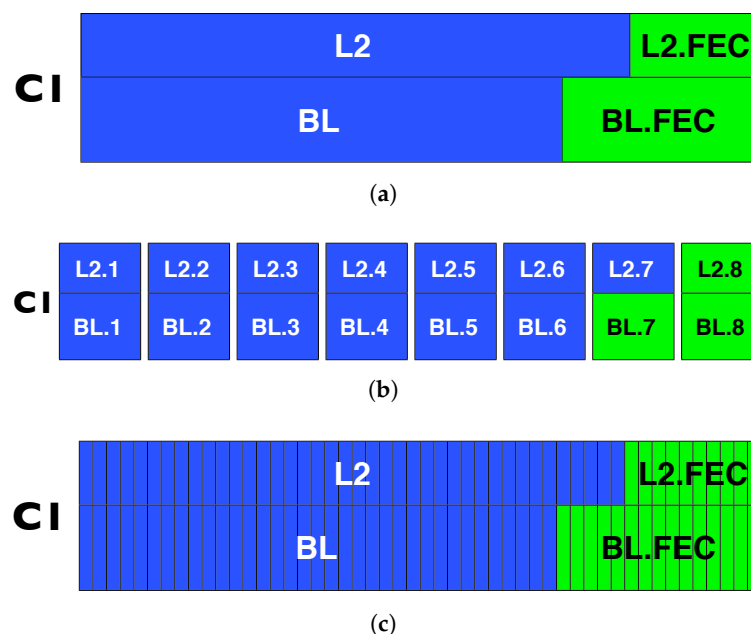


Figure 8. Examples of two layers allocated to class one (C_1) for (a) SVC plus FEC and (b) description-based models plus FEC; (b) illustrates the section structure of MDC utilised by IRP, while (c) illustrates an example of the IRP packetisation of the SVC class in (a).

Streaming Classes—Class Packetisation Evaluation Results

In this section, we illustrate a simple ALD example where two classes are packetised using ROP and IRP and transmitted over a link with a 10% loss rate. We begin by allocating the layers to the classes. In this example, we will use the crew media clip encoded as a two-resolution, five-layer stream.

We allocate the layers to the classes based on resolution, such that two layers (BL and L2) are allocated to C1 and three layers (L3, L4 and L5) allocated to C2. A class composition of Independent Class Composition (ICC) is used. Table 4 illustrates the JSVM output for the first frame.

Table 4. Encoder output for a single frame for the crew video.

Length	LId	TId	QId
18	0	0	0
1424	0	0	0
1577	0	0	1
2186	1	0	0
2461	1	0	1
2388	1	0	2

Table 5 shows the per layer transmission cost for SVC, MDC, ALD with an STF of 6 (shown in Figure 9) and ALD-SC (ALD using a Streaming Class model) with ICI class composition. In our example, we incorporate three FEC sections of the base layer and two FEC sections of layer two to class 1, while also allocating one FEC section of both the base layer and layer two to class two, as illustrated in Figure 10.

Table 5. Transmission cost for SVC, MDC, ALD and ALD-SC.

Layer	SVC	MDC	ALD	Class	ALD-SC
L5	10,054	20,270	12,232	Class 2	12,232
L4	7666	16,216	11,120		
L3	5205	12,162	10,008		
L2	3019	8108	8896	Class 1	4032
BL	1442	4054	7784		

L5.1	L5.2	L5.3	L5.4	L5.5	L5.6	L5.7	L5.8	L5.9	L5.10	L5.11
L4.1	L4.2	L4.3	L4.4	L4.5	L4.6	L4.7	L4.8	L4.9	L4.10	L4.11
L3.1	L3.2	L3.3	L3.4	L3.5	L3.6	L3.7	L3.8	L3.9	L3.10	L3.11
L2.1	L2.2	L2.3	L2.4	L2.5	L2.6	L2.7	L2.8	L2.9	L2.10	L2.11
BL.1	BL.2	BL.3	BL.4	BL.5	BL.6	BL.7	BL.8	BL.9	BL.10	BL.11
Dc - 1	Dc - 2	Dc - 3	Dc - 4	Dc - 5	Dc - 6	Dc - 7	Dc - 8	Dc - 9	Dc - 10	Dc - 11

Figure 9. ALD with five-layer and an STF of 6.

C2	L5.1	L5.2	L5.3	L5.4	L5.5	L5.6	L5.7	L5.8	L5.9	L5.10	L5.11
	L4.1	L4.2	L4.3	L4.4	L4.5	L4.6	L4.7	L4.8	L4.9	L4.10	L4.11
	L3.1	L3.2	L3.3	L3.4	L3.5	L3.6	L3.7	L3.8	L3.9	L3.10	L3.11
C1	L2.1	L2.2	L2.3	L2.4	L2.5	L2.6	L2.7	L2.8	L2.9	L2.10	L2.11
	BL.1	BL.2	BL.3	BL.4	BL.5	BL.6	BL.7	BL.8	BL.9	BL.10	BL.11

Figure 10. Two streaming classes created from ALD with five-layer and an STF of 6. C1 denotes class one and C2 denotes class two.

The reasoning for the FEC section allocation in the lower class is that the $LR_{C_i}^{\max}$ as defined by Equation (12) in Section 4.3 for a 10% packet loss rate is a 17% FEC rate for the highest layer in the class. Because of an STF of 6, eight sections of layer two are required to decode layer two, thus each layer two section equates to 12.5% of layer two. The allocation of a single FEC section for layer two would only equate to 12.5%, which is less than the $LR_{C_1}^{\max}$ of 17% mandated, such that two sections totalling 25% are required to elevate $LR_{C_1}^{\max}$. We view this as an FEC mapping ratio of (3,1), which denotes that the lowest layer in class 1 contains three FEC sections, while the lowest layer in class 2 contains 1 FEC section. This can be generalised to (<number of FEC section in the lowest layer of class 1>, <number of FEC section in the lowest layer of class 2>, ..., <number of FEC section in the lowest layer of class N>). In a streaming class that utilises all FEC sections from the original streaming model, adding all the values in the FEC mapping will equal $N - 1$, where N denotes the total number of layers in the original encoding.

Table 6 presents the percentage of FEC per layer allocated to each class. As we have not varied the transmission cost of ALD-SC with respect to ALD, the cumulative FEC value per layer, e.g., by adding the FEC percentage over both classes for a given layer, are consistent with ALD.

Table 6. Per layer FEC percentage per class for ALD-SC.

Layer	C1	C2
L5	0%	0%
L4	0%	10%
L3	0%	22.24%
L2	25%	12.50%
BL	42.87%	14.29%

As illustrated in Table 5 and Figure 10, in our simple example, we do not increase the transmission cost of ALD and as such do not increase the error resiliency of C2 to the same level of $LR_{C_i}^{\max}$.

Table 7 illustrates the ALD-SC packet sizes and number of packets per packetisation option, ROP and IRP, based on the single frame example in Table 5. It can be seen that, by creating one super section using ROP, the number of packets required per class is noticeably lower than IRP, but understandably the degradation in viewable quality is greater during moments of packet loss. Figure 11 illustrates the percentage of viewable quality for ALD-SC over the duration of the clip for both packetisation schemes, with a loss rate of 10%. Note how a simple reduction in packet payload creates a noticeable increase in the viewable quality. It can also be seen that, due to the incremental levels of FEC per layer within a class, different loss rates per GOP mandate the decodability of different layers per class.

Table 7. ALD-SC Packet sizes and number of packets per packetisation option.

Class	ROP		IRP	
	Packet Size	# of	Packet Size	# of
Class 2	1299	8	800	13
Class 1	1344	3	404	8

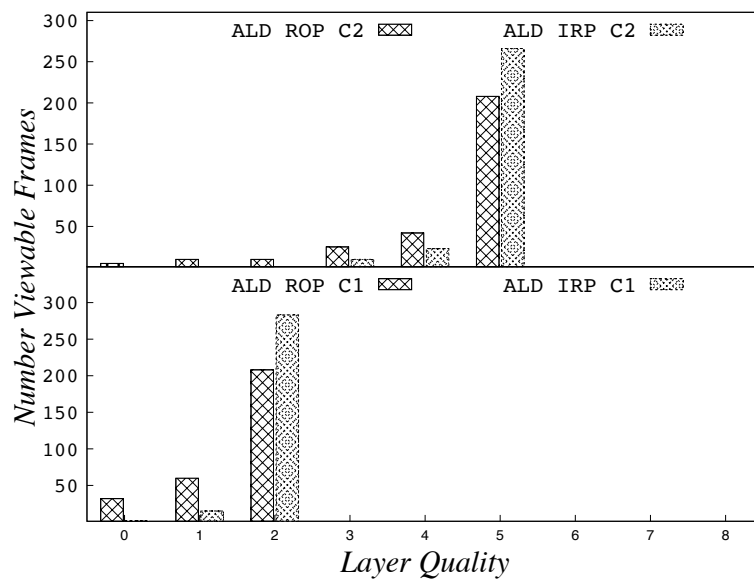


Figure 11. Viewable quality of ALD-SC for both packetisation schemes, ROP and IRP, with a loss rate of 10%.

We extend this example in Appendix A and we provide examples of defining the packet byte-size allocation based on fixed values of 1440 bytes, 1000 bytes and 500 bytes for each class, rather than the defined section size of the underlying description model. Defining the packet byte-size allocation based on fixed byte sizes is consistent with the packet byte-size allocation required for SVC. These values define that all packets per GOP, over all classes, are the same byte size, but not of equal importance (lower class packets are still of higher priority). We also defined a model where a 1440 byte threshold is allocated to the highest class, 1000 byte threshold to the middle class and 500 bytes to the lowest class, thus illustrating byte sizes relative to the underlying byte cost of the respective classes. A final option determines the number of packets for the highest layer using a 1440 byte allocation, and mandated that all lower classes utilise the same number of packets, as a result forcing an optimal packet threshold for the lower classes based on the number of packets for the highest class. The results provided in Appendix A for SVC-SC (SVC using a Streaming Class model) illustrate that, as we increase the GOP value, different levels of allocated FEC can provide continuous levels of maximum stream quality for all packet byte-size allocation schemes illustrated above, thus illustrating both adaptation in FEC allocation and packet byte-size allocation by which viewable quality can be preserved during moments of packet loss. For the examples in this section and Appendix A, a class composition of Independent Class Composition (ICC) is used.

5. Evaluation Framework

The results provided in Appendix A illustrate an example where the Streaming Class design principles are utilised to adapt an SVC stream using SVC-SC (SVC using a Streaming Class model) based on ICC. For the remainder of this section, we present examples of description-based streaming classes, namely MDC and ALD, based on ICI (note we do not increase the transmission cost of MDC or ALD, thus no additional FEC is added to layer 8). It is important to note that the usage of the *Streaming Class Structure* principle for layered or description-based models is interchangeable. The usage of ICC and ICI in our examples is for illustration purposes only, and their allocation to the underlying models demonstrates their usage rather than mandating that ICC is only for SVC-SC and ICI is only for description-based SC.

Prior to presenting our evaluation results, we provide a brief overview of our evaluation framework. Our evaluation is based on the widely-known 10 s *crew* video. The video is encoded using JSVM [77] to eight layers with spatial and quality scalability, using medium grain scalability

(*mgs*), quantizer parameter (*QP*) values for BL to L8 of 34, 28, 33, 30, 28, 35, 32, 30, respectively, and a GOP value of one. We consider three resolutions (QCIF, CIF and 4CIF) with 2, 3, and 3 quality levels, respectively, e.g., two fidelity levels in the lowest resolution and three fidelity levels in each of the higher resolutions, which are allocated to three Streaming Classes based on underlying resolution. QCIF is mapped to C1 (maximum 38.2 dB), CIF is mapped to C2 (maximum 39.3 dB) and 4CIF is mapped to C3 (maximum 38.7 dB).

The transmission of the encoded video is simulated in Network Simulator 2 (*ns-2*) [78] using myEvalSVC [79], an open source tool for evaluating JSVM video traces for SVC. Modifications are made to myEvalSVC scripts to simulate MDC, ALD and SCs. In SVC, each layer per frame is packetised individually, in MDC, each description per frame is packetised separately, while, in ALD, ALD-SC and MDC-SC, each packet contains a segment of each layer per description (using SD). In ALD, this would lead to a segment from every layer per packet, while, in the SC models, each class would be packetised separately, but, within each class, a packet would equate to a segment of each layer.

The simulated network topology is shown in Figure 12 in which we vary the average packet error rate, μ , from 1% to 10% to test the streaming performance of different schemes over lossy links. We use an *ns-2* Errormodel to define a total packet error rate with a uniform distribution, such that the level of loss per frame varies from less than or equal to, to greater than μ , but with total average stream loss equal to μ .

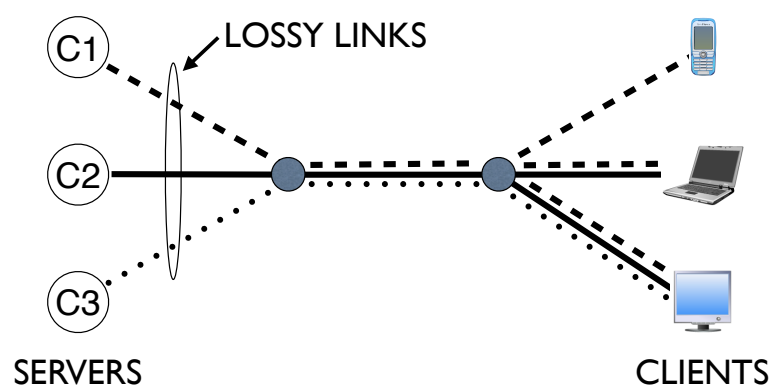


Figure 12. Simulated network topology.

For each of the simulated schemes, sixteen iterations are run to create the *ns-2* output traces, which are analysed to determine the average maximum stream quality per-frame at the client. Each trace is then saved as an achievable quality (AQ) trace file for each streaming scheme. The AQ trace files are utilised to (1) to provide a means of illustrating the transition in frame quality over time and (2) to create the received YUV files, based on the maximum stream quality per frame, from the original YUV files. The metrics we use for evaluating the video quality is Peak Signal to Noise Ratio (PSNR) [80]. PSNR is the simplest and the most widely used video quality evaluation methodology, and is utilised to specify the pixel difference between the transmitted and received video data, on a frame per frame basis (known as full reference), so as to determine a quantifiable value for the variation in viewable quality. In this evaluation, we utilise PSNR to evaluate the effects of packet loss on viewable quality.

6. Simulation Results for SVC, MDC and ALD without Using the SC Framework

The following results are provided as an example of the SVC, MDC and ALD evaluation determined so far and provide a base case comparison to the SC result shown later in this section. Figure 13a plots the Y-PSNR (Y-PSNR is the measured PSNR for the Y-component of YUV) values versus the percentage of datagram loss over the communication link for SVC, MDC and ALD when the user is streaming the highest video quality (4CIF). In this section, SVC is shown for comparison purposes only. The results indicate that ALD shows the best performance followed by MDC and

then SVC. Typically, MDC is better than SVC due to the included FEC. The further improvements achieved by ALD are due to the increase in the number of descriptions, reduction in the byte-allocation per description section and SD. SD disperses the loss impact over several sections instead of a single datagram loss affecting only one layer (SVC) or one description (MDC).

Figure 13b confirms these results by showing the number of frames viewed at every quality level for the crew video at a datagram loss rate of 10% when SVC, MDC and ALD are used. This figure demonstrates the severe impact of packet loss on SVC performance, where approximately 40 frames were completely lost due to the loss of the base layer. Additionally, SVC shows frequent quality-level shifts where each layer is viewed between 17 and 47 times. On the contrary, users using MDC video enjoyed a better streaming experience where more than 80% of the frames are viewed between qualities 3 and 6. Using ALD further improves the streaming experience where more than 80% of the video is shown at the highest resolution (4CIF).

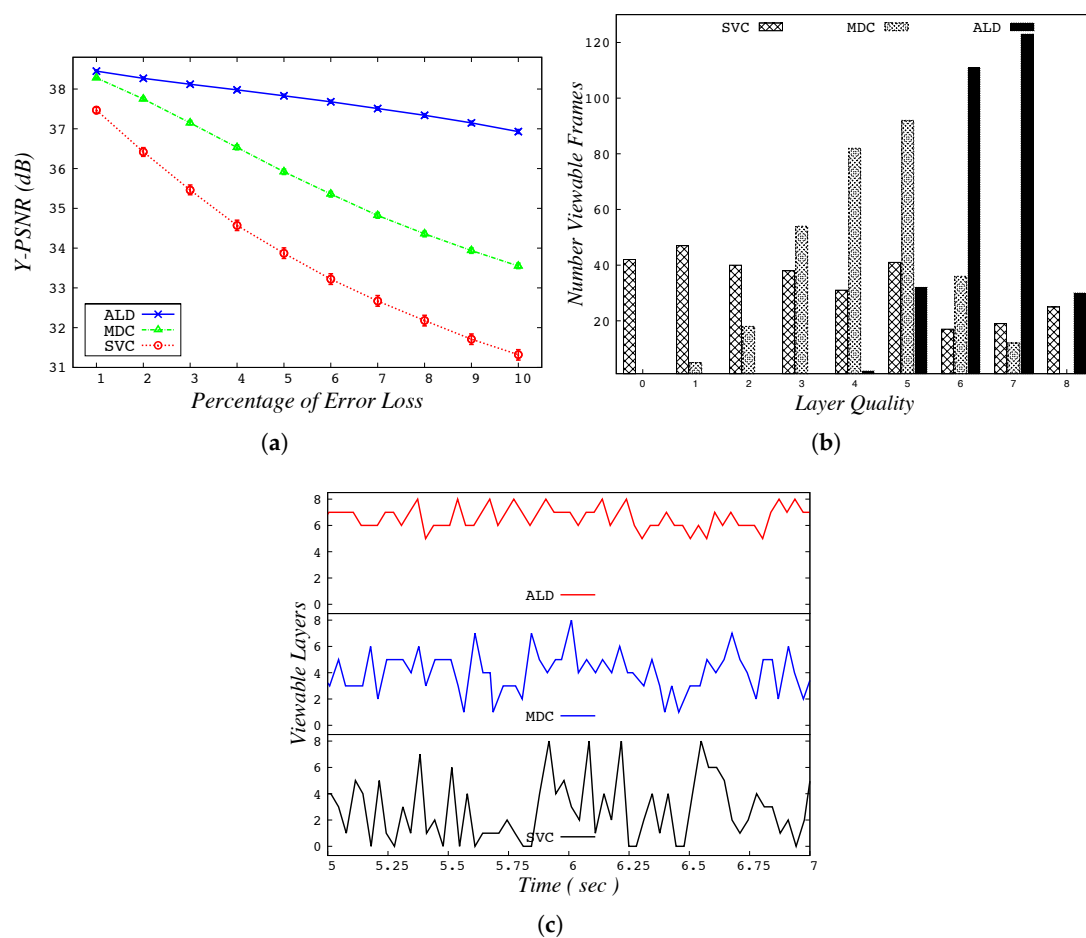


Figure 13. Performance of scalable video encoding over lossy links. (a) versus packet loss ratio; (b) viewable video quality at 10% loss; (c) 2 s sample of viewable quality transitions.

Figure 13c takes the frames viewed at every quality level, for each of the models, and plots a two-second sample of the frequency of layer switching. This result again illustrates the high variation in quality for SVC, minor increase in quality for MDC and consistency of quality in the higher layers for ADL.

More importantly, these improvements are attained at a lower transmission byte-cost as shown in Table 8, which also shows the relative transmission cost compared to SVC. These savings in byte cost are made possible thanks to the STF component of ALD. In the results shown, the value of STF

is 6 according to the developed optimisation framework shown in [56]. Hence, a total of fourteen descriptions are required to stream the video in ALD at the highest quality. The main drawback of section thinning in ALD is the higher transmission cost when low quality video is requested. As we shall show, this problem is eliminated by using SCs.

Table 8. Transmission byte-cost as for the highest quality.

Scheme	SVC	MDC	ALD
Transmission (bytes)	9,483,746	19,334,064	12,075,882
Datagrams	7687	14,248	10,108
Value compared to SVC	100%	≈204%	≈128%

7. SC Performance Evaluation

First, we present the transmission cost in bytes for ALD and the corresponding savings per class when ALD-SC (ALD using a Streaming Class model) is used. Figure 14a plots the transmission byte-cost for the crew video versus the ALD-STF value for distinct video qualities. Note that the eight ALD description encoding, $STF = 0$, degenerates to MDC, thus Figure 14a represents both ALD and MDC. Figure 14a illustrates the aforementioned limitation of ALD showing the increase of the transmission cost of streaming low quality as STF increases. On the contrary, Figure 14a also shows that, as STF increases, the transmission cost of streaming high quality video decreases. Figure 14b plots the transmission cost of SCs versus the number of ALD descriptions. It can be seen that, by using ALD-SC, there is a significant drop in transmission cost and that the transmission cost of ALD-SC always decreases as STF increases.

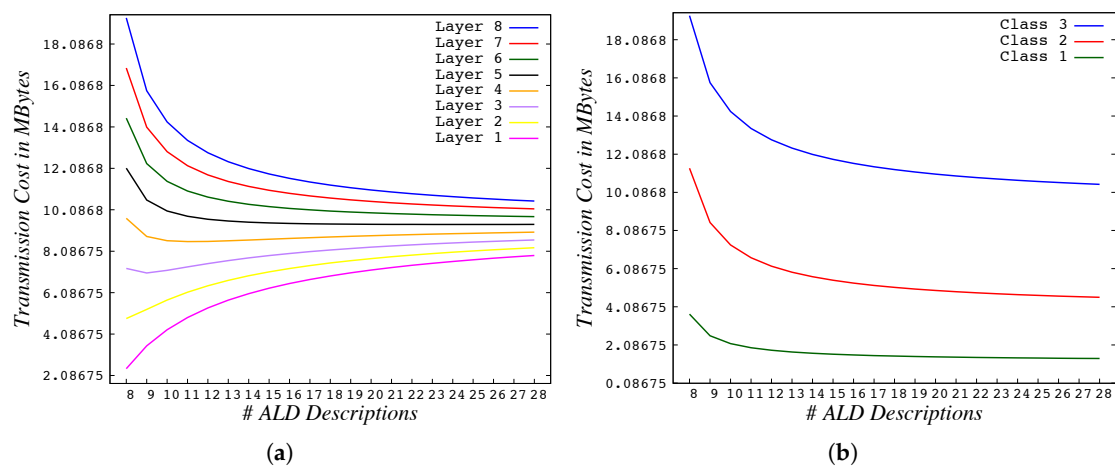


Figure 14. Transmission cost of the crew media clip for (a) each layer; and (b) each streaming class, as the STF value and associated number of ALD descriptions increase.

Table 9 illustrates the per quality cumulative transmission cost for SVC, MDC, ALD ($STF = 6$) and the respective SC schemes. The number in the brackets denotes the number of transmission units (layers for SVC, descriptions for MDC and ALD, and classes for the SCs) required to decode the target quality. It can be seen that, due to the packetisation of MDC-SC, there is a reduction in the transmission cost of layer 2 and layer 5, with respect to MDC. It is important to note that these figures are content and encoding specific. Minor modifications to the quantisation parameters (QP) in JSVM and subsequent decoding quality can reduce the increase costs of streaming classes. For ALD-SC, the transmission cost values show a marked reduction in the transmission overhead for the quality layers within the lower classes, with respect to all other models. ALD-SC class allocation for layers six and

seven mandate a minor increase in transmission cost of ALD-SC relative to ALD, due to the allocation of lower layer FEC data within C3. However, it is important to note that such minor increases in transmission cost, for both MDC-SC and ALD-SC, is accompanied by an improvement in achievable quality due to layer grouping. Note: as each class is composed of layer data and incremental levels of FEC, only a subset of packets is required to decode the lower layers per class, thus implementing a simple feedback mechanism for deterring the number of packets required per layer, per class. Lower layer streaming per class could be performed, thus further reducing decoding complexity at the device. Overall transmission cost would not be decreased, as all packets per class would still be transmitted, whereby only a subset of packets would be received. Table 9 also shows the number of transmitted datagrams for each SC scheme assuming a maximum transmission unit of 1500 bytes. It is important to note that transmitting more datagrams implies an increased transmission overhead for lower layers (not shown).

Table 9. Transmission byte cost for Different Encoding Schemes.

Layer	SVC	MDC	ALD	MDC-SC (ROP)	Dg	ALD-SC (ROP)	Dg	ALD-SC (IRP)	Dg
8	9,483,746 (8)	19,334,064 (8)	12,075,882 (14)						
7	7,419,782 (7)	16,917,306 (7)	11,213,319 (13)	19,347,198 (3)	13,892	12,077,528 (3)	8825	12,077,528 (3)	13,558
6	5,640,092 (6)	14,500,548 (6)	10,350,756 (12)						
5	3,931,226 (5)	12,083,790 (5)	9,488,193 (11)						
4	2,993,946 (4)	9,667,032 (4)	8,625,630 (10)	9,726,545 (2)	7072	5,663,237 (2)	4215	5,663,237 (2)	6348
3	2,044,662 (3)	7,250,274 (3)	7,763,067 (9)						
2	1,232,574 (2)	4,833,516 (2)	6900,504 (8)						
BL	617,526 (1)	2,416,758 (1)	6,037,941 (7)	2,776,152 (1)	2090	1,651,653 (1)	1281	1,651,653 (1)	2700

In the following, we compare the streaming quality performance of MDC, MDC-SC, and ALD-SC by showing PSNR and viewable frames different quality levels when ROP and IRP are used. As illustrated in Section 4.5, with larger STF values, ALD-SC requires a greater number of FEC sections in the lower classes to accommodate greater levels of packet loss. Thus, in this comparison, an FEC mapping ratio (see Section 4.5) of 2, 3, 2 is used for MDC while ALD-SC has an FEC mapping of 3, 3, 1. The mapping ratio of ALD-SC is larger in C1 due to the smaller byte-allocation per lower layer section, the increased impact of datagram loss on achievable quality and the $LR_{C_i}^{\max}$ as defined by Equation (12) in Section 4.3, thus mandating a higher level of FEC for C1. MDC has larger levels of inherent FEC per section with which to combat packet loss, thus a lower level of allocated FEC is required for C1.

Figure 15a plots the number of viewable frames per quality level for each of the two SC classes (MDC-SC and ALD-SC) at 10% datagram loss rate using ROP for the streaming classes, as well as the respective highest layer in MDC: Layer 2 (L2) in C1, Layer 5 (L5) in C2, and Layer 8 (L8) in C3). SVC has been removed from all subsequent plots as we are now comparing description based models. SVC-SC is presented in Appendix A. The three subfigures may be considered a representation for users with different bandwidth availability. Clearly, the figure shows that MDC-SC has the best performance for all user types followed by ALD-SC then MDC. Noting that both MDC and MDC-SC stream identical data (same number of sections) when the highest quality is requested (top most figure), the performance gain for MDC-SC over MDC is interpreted by the positive impact of SD over the defined super sections. For the same highest quality, the success of ALD in decoding more frames at the highest quality (layer 8) is due to SD packetisation. For limited bandwidth users (bottom subfigure), the included FEC sections are considered the key reason for MDC-SC and ALD-SC in achieving a higher video quality in comparison to MDC. A similar performance is noticed in the middle figure (intermediate bandwidth availability).

Figure 15b highlights the average Y-PSNR values versus the packet loss ratio for the same considered schemes. Note that the SC grouping for ALD-SC and MDC-SC has provided near consistent quality for each iteration of the simulation, thus mandating a near un-viewable range of confidence

interval error bars. Clearly, Figure 15b is consistent with the results in Figure 15a since MDC-SC achieves the highest PSNR followed by ALD-SC then MDC. The figure shows a 2 dB difference in the PSNR between MDC-SC and ALD-SC when the highest quality is streamed at 10% loss. For the same quality and loss ratio, a larger performance gap of 4 dB exists between MDC-SC and MDC, due to the SD component in SCs. This PSNR gap is much smaller for the intermediate and limited bandwidth cases (middle and bottom figures). It can also be seen that, by utilising ALD-SC with ROP, so as to reduce transmission cost for lower classes, the PSNR values of ALD-SC have dropped when compared with ALD from Figure 13a.

Figure 15c illustrates a two-second sample of the frequency of layer switching for each of models. Note the high impact of packet loss on the variation of quality for each model, which occurs with ROP packetisation.

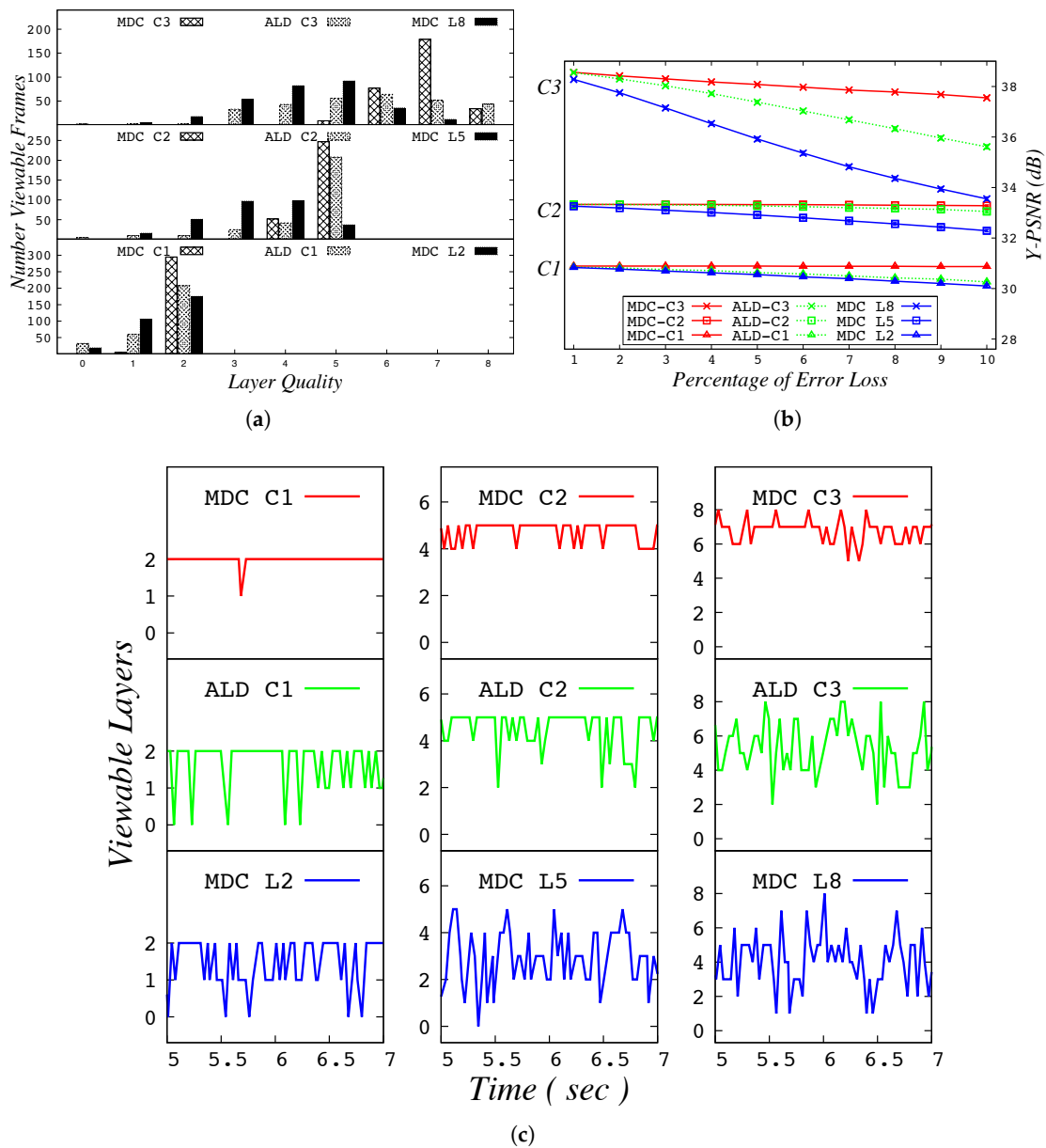


Figure 15. Performance evaluation of considered schemes using ROP for SCs. (a) video quality for different schemes at 10% loss; (b) average Y-PSNR values for different schemes with 95% confidence interval results; (c) 2 s sample of viewable quality transitions for MDC (layers L2, L5 and L8) and for each of the classes of ALD-SC and MDC-SC (maximum quality per class equating to layers 2, 5 and 8).

In practice, PSNR for a selected resolution is analysed against the same resolution. However, in our figures, this would have created increased PSNR values for all lower classes, where the goal of the PSNR figures is to highlight the consistency of the quality, per class, as datagram loss increases. Thus, in our evaluation, we calculated the PSNR values for each model per class, irrespective of original resolution, by comparing the modified YUV file against the original YUV file with the highest quality and resolution, e.g., Layer 8. This creates reduced PSNR values for the lower classes, but provides PSNR values that are consistent with the achievable quality of each layer.

Table 10 outlines a summary of the mean layer value viewed over the duration of the clip, i.e., over all 300 frames, and the maximum, \max_{qual} , and minimum, \min_{qual} , layer value viewed, thus illustrating the variation in quality that occurred during decoding. We will comment on each one of the streaming models separately. For MDC-SC, we note that each additional class received can increase mean layer quality to each respective resolution. For ALD-SC, we note that the non-decodable frames in C1 cascade over all higher classes, thus reducing overall viewable quality for these specific frames. However, for MDC, the quality is dependent on the number of descriptions received without data loss, thus MDC L5 has a minimum decodable layer quality of layer 1, while MDC L8 is unable to decode some frames. The selective packetisation utilised in MDC-SD illustrates the increase in achievable quality with no increase in transmission cost with respect to MDC.

Table 10. Example of the mean layer value and the variation that occurs between maximum, \max_{qual} , and minimum, \min_{qual} , layer viewable quality over the duration of the clip.

	Mean	\min_{qual}	\max_{qual}
MDC C1	1	1	2
MDC C2	4	3	5
MDC C3	6	4	8
ALD C1	1	0	2
ALD C2	4	0	5
ALD C3	5	0	8
MDC L2	1	0	2
MDC L5	3	1	5
MDC L8	4	0	8

On using IRP, the performance gap between ALD-SC and MDC-SC shrinks to 1 dB for the highest quality at 10% loss. Additionally, the PSNR performance gap becomes insignificant for both low and intermediate quality levels. Figure 16a–c respectively show the same performance metrics as Figure 15a–c but for IRP packetisation. This performance gain is attained due to distributing the error impact over a larger number of smaller packets. However, these packets also introduce an additional transmission cost of extra packet headers belonging to lower layers. For the crew video, this additional overhead can be estimated as a 3% increase in the total transmission cost in IRP in comparison to ROP (assuming a 60-byte header in a 1500 byte packet). In conclusion, the additional overhead of MDC-SC is considered useful only for users having abundant bandwidth and lossy links. In case of limited or low bandwidth, ALD-SC performs similarly to MDC-SC but with a much lower overhead.

Table 11 outlines a summary of the mean layer value viewed over the duration of the clip, i.e., over all 300 frames, and the maximum, \max_{qual} , and minimum, \min_{qual} , layer value viewed, thus illustrating the variation in quality that occurred during decoding. We will comment on each one of the streaming models separately. For MDC-SC, we note similar results to Table 10 but an increase in the minimum quality layer from layer 4 to layer 5 for C3. For ALD-SC, we note that, due to IRP as well as the underlying ICI class structure, there is an increase in the minimum quality layer from a non-decodable frame to layer 3 for C2 and C3. The results for MDC remain the same as for Table 10, as MDC does not use either ICI or IRP.

Table 11. Example of the mean layer value and the variation that occurs between maximum, \max_{qual} , and minimum, \min_{qual} , layer viewable quality over the duration of the clip.

	Mean	\min_{qual}	\max_{qual}
MDC C1	1	1	2
MDC C2	4	3	5
MDC C3	6	5	8
ALD C1	1	0	2
ALD C2	4	3	5
ALD C3	6	3	8
MDC L2	1	0	2
MDC L5	3	1	5
MDC L8	4	0	8

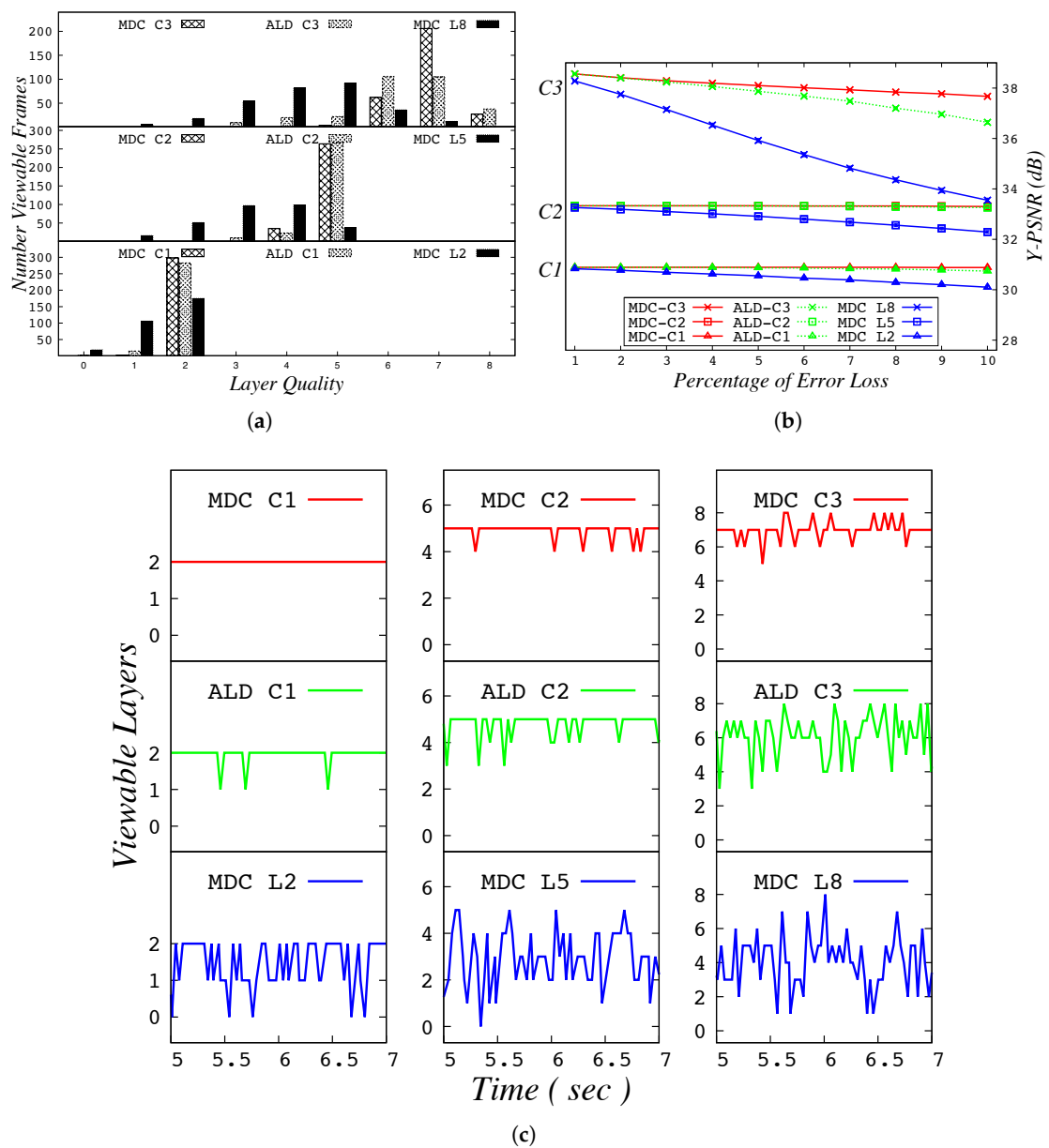


Figure 16. Performance evaluation of considered schemes using IRP for SCs. (a) video quality for different schemes at 10% loss; (b) average Y-PSNR values for different schemes with 95% confidence interval results; (c) 2 s sample of viewable quality transitions.

8. Discussion

As we have seen from the evaluation results from Sections 4.5 and 7 and Appendix A, the design principles of SC increase the viewable quality of the underlying streaming models. Section 7 presents our evaluated results for description-based SC streaming (MDC-SC and ALD-SC) in comparison to defined layer values in MDC. Section 4.5 illustrates a simple example of the variation in quality of ALD dependent on the packetisation options ROP and IRP, while Appendix A extends the granularity of the packetisation by showing the variation in quality over varying GOP values for an SVC-SC model. Each of our design principles provides a means of adapting elements of the stream flow to suit network and user requirements, as well as illustrating the interaction that can occur between the elements. For layered schemes such as SVC, this mandates a minor increase in transmission cost, relative to the levels of error resiliency allocated, while, for description-based schemes such as MDC and ALD, the transmission cost can be reduced and optimised to suit achievable quality. Our evaluated examples provide a sample of the possible configuration options that can be selected by using our streaming class framework, and even these samples illustrate the gains that can be made in viewable quality.

What can be inferred from our results:

- That increases in transmission cost as mandated by MDC do not always provide for stability in achievable quality.
- That strategic choices made during the various elements of the stream flow can increase viewable quality.
- That the inherent complexity in the number of options offered by SC may overly complicate the decision-making process.
- That a larger test sample offered by a complete real-world implementation of SC is required to fully appreciate the operations and interdependency between the elements of the network flow and our design principles.
- That no one group of SC options is sufficient to accommodate for all streaming models and transmission mediums, and the issues that can occur during streaming over the network.

9. Conclusions

In this paper, Streaming Classes (SC) is proposed as a new approach for the transmission of layered and description-based adaptive streaming. We present five design principles by which high levels of viewable quality can be maintained over the duration of the entire media clip. We propose that data belonging to different quality layers are grouped together to form streaming classes, based on user, network or scalable dimensionality requirements, by which transmission byte-cost and achievable quality can be managed and improved respectively. Evaluation results for streaming class extension of existing video models (SVC-SC, MDC-SC and ALD-SC) show significant performance improvements such as consistent high levels of quality for users with varying resource availability, as well as a reduction in transmission cost for devices requiring lower layer description-based decoding.

In this research, we learned that maintaining the consistency of achievable quality comes at a cost (prioritisation, increased transmission cost and complexity), but the benefits of consistent achievable quality over the duration of the media stream, for heterogeneous devices, irrespective of layer requirement, far outweigh these underlying costs. We also learned that loss naturally occurs in the transmission network and simply pushing video data into the network without consideration of said loss, the needs of users, or the optimisation of the stream flow is neither beneficial to the transmission costs of network providers or their clients, or mandates consistent of achievable quality for users.

Acknowledgments: This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grants 13/IA/1892 and 10RFP/CMS2952.

Author Contributions: All authors conceived and designed the experiments; Jason J. Quinlan performed the experiments; all authors analyzed the data; Jason J. Quinlan wrote the paper, with considerable assistance from Ahmed H. Zahran and Cormac J. Sreenan.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Additional Evaluation

In this Appendix, we extend the example in Section 4.5 and provide examples of defining the packet byte-size allocation based on fixed values of 1440 bytes, 1000 bytes and 500 bytes for each class, rather than the defined section size of the underlying description model. Defining the packet byte-size allocation based on fixed byte sizes is consistent with the packet byte-size allocation required for SVC, thus we shall view these evaluation results as SVC-SC (SVC using a Streaming Class model). These values define that all packets per GOP, over all classes, are the same byte size, but not of equal importance (lower class packets are still of higher priority). We also defined a model where a 1440 bytes threshold is allocated to the highest class, a 1000 bytes threshold to the middle class and a 500 bytes threshold to the lowest class, thus illustrating bytes sizes relative to the underlying byte cost of the respective classes. A final option determines the number of packets for the highest layer using a 1440 bytes allocation, and mandated that all lower classes utilise the same number of packets, as a result forcing an optimal packet threshold for the lower classes based on the number of packets for the highest class. Finally, we illustrate results for these packet byte-size allocation models over one loss rate of 10% and over four GOP values, e.g., GOP-1, GOP-8, GOP-16 and GOP-32.

The focus of this appendix is to investigate the error resilience of SC under different SC structures and packetisation options. Of the five Streaming class design principles, as presented in Section 4, the allocation rates of the principles of *Error Resiliency* and *Class Packetisation and Granularity of Packet Data Byte-size* are the subject of this Appendix. For the remainder of the design principles, in our evaluated results, we mandate a *Class Packet Loss Rate* of 10%, a *Layer Allocation and Hierarchy* of three classes based on an eight layer SVC encoding, with class composition based on resolution and a *Streaming Class Structure* of Independent Class Composition (ICC) is used.

We begin by defining the values for the FEC allocation options $LR_{\max} (\lceil \mu + \sqrt{\mu} \rceil)$ and $LR_{C_i}^{\max}$ (see Equation (11) in Section 4.3) based on the 10% loss rate, as illustrated in Table A1. Table A1 also illustrates the *Layer Allocation and Hierarchy* of the three classes. In this example, we extend the LR_{\max} and $LR_{C_i}^{\max}$ values over all classes, thus increasing the overall transmission cost of SVC-SC. We use the same initial LR_{\max} and $LR_{C_i}^{\max}$ allocation for all classes, e.g., 14% and 17%, respectively. Note how the level of FEC is based on the index of the layer in the underlying class, with higher layers receiving lower levels of FEC and lower layers receive allocation rates based on the preceding higher layer plus the standard deviation of the previous higher layer, i.e., the FEC allocation rates of layer 7 are based on the FEC rates of layer 8 plus the square root of the FEC rates of layer 8. Using this allocation rate of FEC, it can be seen that, while C1 is the highest priority class, it contains the lowest levels of FEC, due to only two layers being allocated to this class.

Table A1. LR_{\max} and $LR_{C_i}^{\max}$ SVC-SC packet byte-size allocation per layer, and per class, for 10% packet loss rates.

Layer	LR_{\max}	$LR_{C_i}^{\max}$	Class
L8	14%	17%	Class 3
L7	18%	22%	
L6	23%	27%	
L5	14%	17%	Class 2
L4	18%	22%	
L3	23%	27%	
L2	14%	17%	Class 1
BL	18%	22%	

In the next option, illustrated in Table A2, we apply the initial LR_{\max} and $LR_{C_i}^{\max}$ allocation for all classes, e.g., 14% and 17% respectively, to the highest layer only, layer 8. We then view the stream based on layer dependency rather than class structure, and extend increasing levels of LR_{\max} and $LR_{C_i}^{\max}$ over all layers. This provides a direct correlation between the level of FEC and the priority of the underlying layer. In this option, the FEC rates of C3 are unchanged, while the rates of C1 and C2 have increased. While the options shown in Tables A1 and A2 can be viewed as the minimum (best-case) and maximum (worst-case) FEC allocation rates for our schemes, Table A2 can also be viewed as a outline of the possible LR_{\max} and $LR_{C_i}^{\max}$ rates that can be used. An example of this would be to increase the allocation rate for C1 from 17% and 22% to 22% and 27% for LR_{\max} and from 14% and 18% to 18% and 23% for $LR_{C_i}^{\max}$, thus providing a means of adapting the FEC allocation even further so as to provide increased adaptation to loss or to degradation in viewable quality.

Table A2. Based on initial loss rate, maximum Loss Rate LR_{\max} and $LR_{C_i}^{\max}$ SVC-SC allocation per layer, and per class, for a 10% packet loss rate.

SVC-SC 10%			
Layer	LR_{\max}	$LR_{C_i}^{\max}$	Class
L8	14%	17%	Class 3
L7	18%	22%	
L6	23%	27%	
L5	28%	33%	Class 2
L4	34%	39%	
L3	40%	46%	
L2	47%	53%	Class 1
BL	54%	61%	

Appendix A.1. Varying Packet Size for a GOP of One

Table A3 illustrates the GOP1 transmission cost for SVC, MDC, ALD (STF = 6) and SVC-SC for the FEC allocation rates LR_{\max} and $LR_{C_i}^{\max}$ as per Table A1. Figure A1 plots the transmission cost for the SVC-SC classes C1, C2 and C3 for LR_{\max} and $LR_{C_i}^{\max}$ with packet loss rates from 0% to 10% based on the FEC rates of Table A1. Note the slight increase in transmission cost for $LR_{C_i}^{\max}$ over LR_{\max} . Note that the transmission cost with a 0% packet loss rates denotes the standard SVC transmission cost of each of the highest layers per class.

Table A3. Transmission costs per layer for SVC, MDC, ALD, and per class for both Forward Error Correction (FEC) allocation options for SVC-SC, for a packet loss rate of 10%. A GOP value of one and the FEC allocation rates as per Table A1 are implemented for these results.

Layer	Existing			Class	SVC-SC 10%	
	SVC	MDC	ALD		LR_{\max}	$LR_{C_i}^{\max}$
L8	9,483,746	19,334,064	12,075,882	Class 3	11,176,159	11,519,753
L7	7,419,782	16,917,306	11,213,319			
L6	5,640,092	14,500,548	10,350,756			
L5	3,931,226	12,083,790	9,488,193	Class 2	4,619,070	4,760,894
L4	2,993,946	9,667,032	8,625,630			
L3	2,044,662	7,250,274	7,763,067			
L2	1,232,574	4,833,516	6,900,504	Class 1	1,430,363	1,473,636
BL	617,526	2,416,758	6,037,941			

From this point forward, we are going to use FEC to denote LR_{\max} , as this is our basic level of FEC and we are going to use FEC_{\max} to denote $LR_{C_i}^{\max}$, as this is our maximum level of FEC. This is purely to make the legends and text in the plots clearer and more legible. This would change the notation in Figure A1 from LR_{\max} and $LR_{C_i}^{\max}$ to the notation FEC and FEC_{\max} as used in Figure A2. Note that the plots lines in Figures A1 and A2 are identical as only the notation has changed.

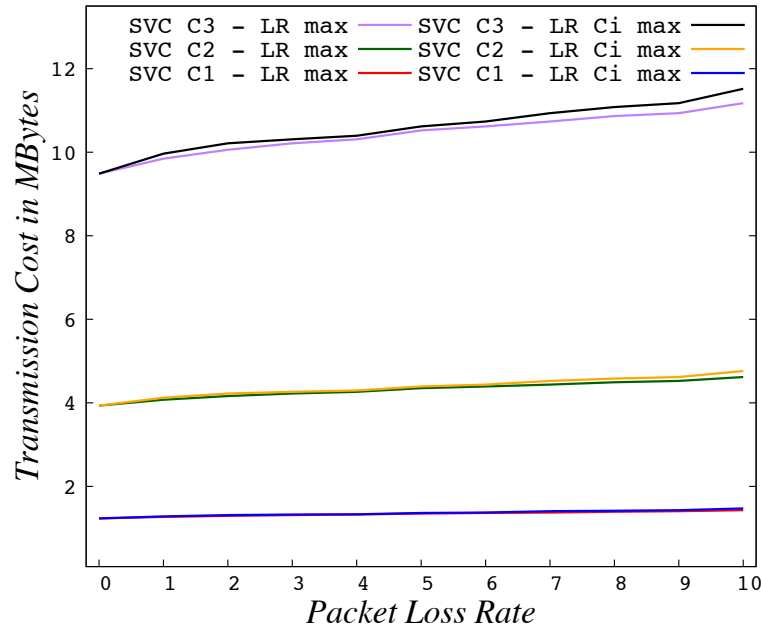


Figure A1. Transmission cost for the SVC-SC classes C1, C2 and C3 for both LR_{\max} and $LR_{C_i}^{\max}$ with packet loss rates from 0 to 10%.

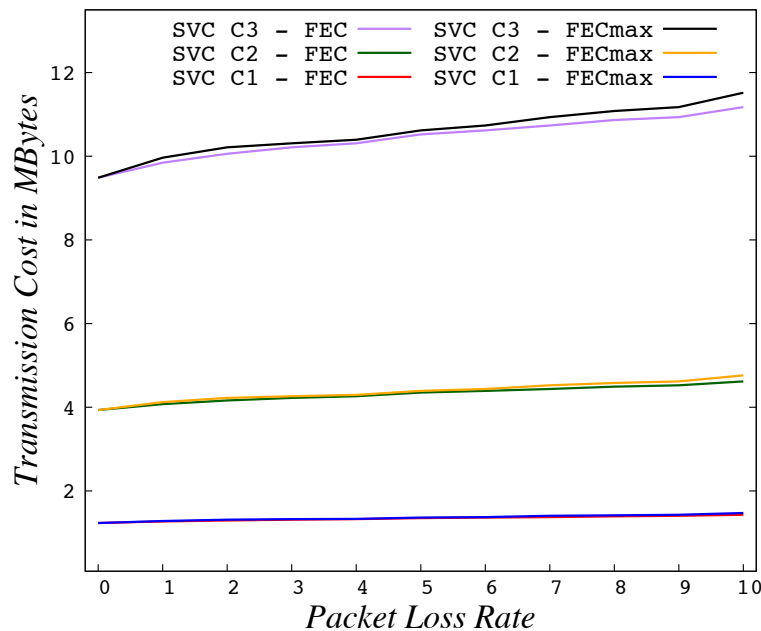


Figure A2. Transmission cost for the SVC-SC classes C1, C2 and C3 for both FEC and FEC_{\max} with packet loss rates from 0 to 10%.

Evaluation results are provided for both FEC and FEC_{\max} and for five different packet byte-size allocation schemes, where each of the classes is defined based on the same packet byte-size allocation, e.g., three different Packet thresholds, $Packet_{thres}$, are simulated: 1440 bytes, 1000 bytes and 500 bytes. We also evaluate where each of the classes is defined based on a different $Packet_{thres}$, e.g., 500 bytes

for C1, 1000 bytes for C2 and 1440 bytes for C3, and we call this *multi*, e.g., multiple different class byte-allocations per stream. Additionally, we consider another packetisation scheme in which we evaluate based on the number of packets transmitted. Rather than define a different $Packet_{thres}$ for each class, we shall define the same number of packets for each class, based on the number of packets defined for the highest class. Thus, the byte-allocation per packet is dependent on the bitrate per class. We call this scheme ' $Packet_{equal}$ '. ' $Packet_{equal}$ ' similar to *multi* in that different $Packet_{thres}$ are defined for each of the classes, but the $Packet_{thres}$ of ' $Packet_{equal}$ ' is dynamically allocated dependent on the number of packets created for the highest class, as apposed to the static $Packet_{thres}$ allocation of *multi*. For extremely low quality streaming, situations may occur where ' $Packet_{equal}$ ' may mandate that the byte allocation of the packets of the lowest classes contain only a small number of bytes. This would occur when the quality of the encoded resolutions and their underlying bitrates are vastly different, especially with reference to the highest and lowest quality levels, e.g., Full HD (1920×1080) at layer 8 and QCIF (176×144) at the base layer, but this allocation of resolution during encoding would rarely, if ever, be selected.

Table A4 shows for each of the packet byte-size allocation schemes: number of packets sent per class, and per stream, for a defined loss rate, 10%, the header cost of these packets, the packets lost per class, the loss rate determined per class and the maximum per frame loss rate. Note that the percentage values are rounded up.

It can be seen that there is little variation in the overall loss rates per class, relative to the loss rate experienced, thus illustrating that the loss rates are equally spread over all classes. The 500 byte packet threshold has the largest number of packets transmitted followed by $Packet_{equal}$, and both of these options have the lowest maximum loss rate per frame for all five schemes—giving us our first indication that lower byte cost, or a higher number of packets, can spread the effects of network loss and improve viewable quality.

Table A4. FEC and FEC_{max} models for all four $Packet_{thres}$, as well as $Packet_{equal}$. Values are illustrated for packets transmitted, additional packet header cost (60 bytes per packet for header information) and individual packets sent, packets received, loss rate per class and maximum loss rate per frame for SVC-SC, for an overall packet loss rate of 10% and a GOP of one.

SVC-SC 10% Different Packet Loss Rates Per Class, with a GOP of 1										
	FEC					FEC_{max}				
Byte Size	1440	1000	500	Multi	$Packet_{equal}$	1440	1000	500	Multi	$Packet_{equal}$
# Packets	8215	11,594	22,776	11,034	14,082	8469	11,969	23,487	11,379	14,532
Header Cost	492,900	695,640	1,366,560	662,040	844,920	508,140	718,140	1,409,220	682,740	871,920
Packets Sent										
C1	1169	1158	3000	3000	4694	1194	1609	3096	3096	4,844
C2	2352	3340	6516	3340	4694	2431	3439	6725	3439	4,844
C3	4694	6696	13,260	4694	4694	4844	6921	13,666	4844	4,844
Packets Received										
C1	1038	1401	2733	2686	4247	1068	1445	2786	2786	4374
C2	2101	3019	5816	3011	4204	2196	3107	6082	3084	4338
C3	4256	6001	11,943	4225	4224	4360	6211	12,271	4361	4365
Loss Rate (LR)										
C1	11%	10%	9%	10%	10%	11%	10%	10%	10%	10%
C2	11%	10%	11%	10%	10%	10%	10%	10%	10%	10%
C3	9%	10%	10%	10%	10%	10%	10%	10%	10%	10%
Max per Frame LR										
C1	100%	80%	44%	50%	45%	75%	75%	50%	50%	33%
C2	50%	40%	30%	56%	36%	44%	44%	30%	42%	46%
C3	33%	38%	22%	46%	42%	50%	35%	25%	42%	43%

For Figure A3, it can be seen that the $Packet_{thres}$ of '500', 'multi' and $Packet_{equal}$, called 'equalP' in the plot, provide the highest number of viewable frames for layer 8, with the lowest number of non decodable frames (un-viewable frames). However, the viewable quality across all layers and models is very bad overall. Figure A4 does prove that the minor increase in FEC provided by FEC_{max} is very beneficial to the overall quality, with marked reductions in all the lower layers and noticeable increases in the highest layer, but the noticeable levels of non decodable frames (un-viewable frames), approximately 20% for $Packet_{thres}$ of '500' and 'multi', and 13% for 'equalP' is still too high.

One item to note across both FEC models is that the consistency of quality for layers 3 and 6 (lowest layers for C2 and C3) is very similar, with this consistency also evident in layers 4 and 7 (mid layers for C2 and C3) for FEC_{max} only. We believe this illustrates that lower levels of FEC will provide consistent quality for different class streaming models, but an FEC level higher than FEC_{max} will be required.

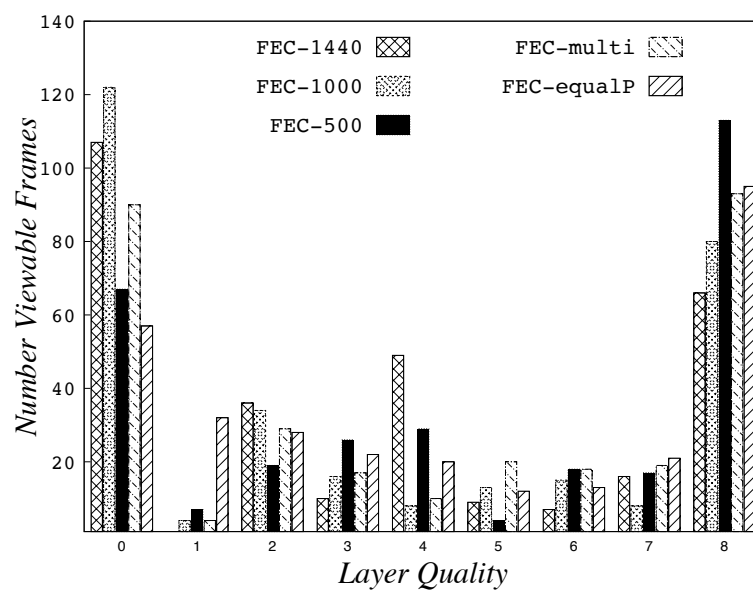


Figure A3. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{Packet_{thres}}$, at a packet loss rate of 10%.

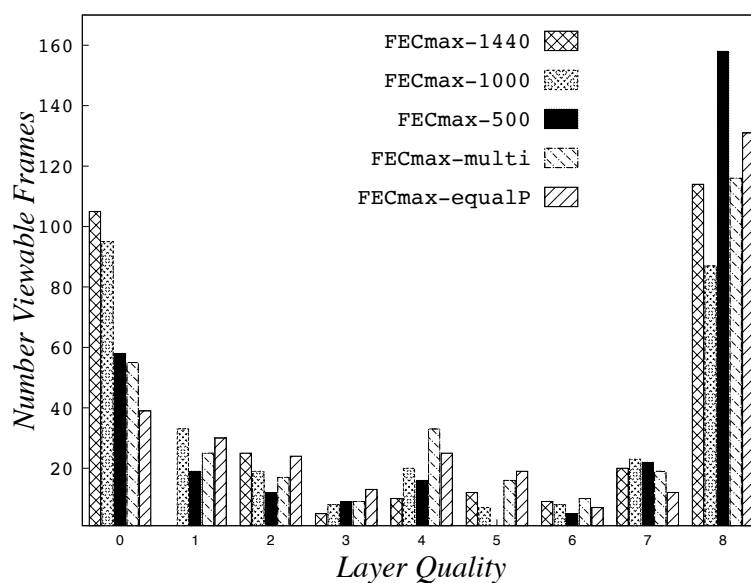


Figure A4. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{max}Packet_{thres}$, at a packet loss rate of 10%.

Figures A5 and A6 illustrate a two-second example of variation in viewable quality for all packet byte-size allocation schemes with reference to Figures A3 and A4, respectively. Note that all schemes fail to provide consistency of quality. Thus, it can be determined that, for this GOP value and these FEC levels, consistency of quality is unavailable and degradation of quality will occur. All of the models show wide variation in viewable quality over time, while ‘500’ for FEC_{max} does show some increase in quality, but this is limited to only a subset of the viewable frames and would only provide limited increases in perceived quality.

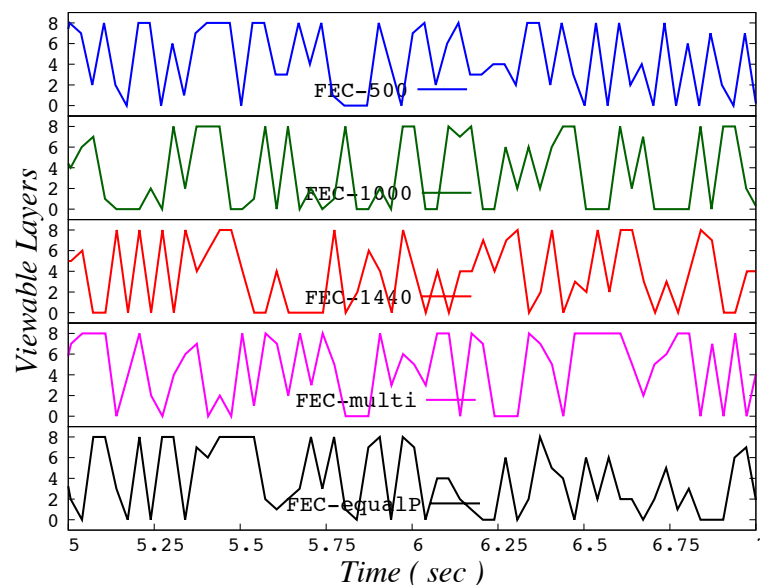


Figure A5. A two-second example of variation in viewable quality for all FEC models, at a packet loss rate of 10%.

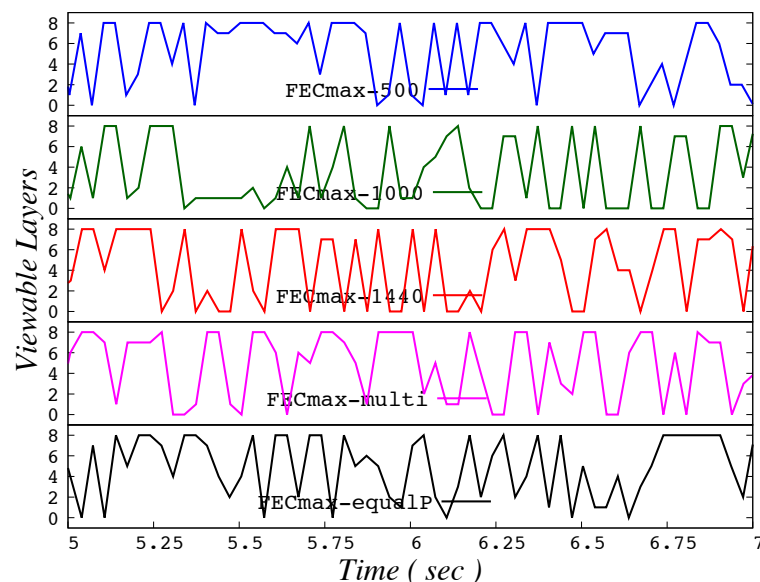


Figure A6. A two-second example of variation in viewable quality for all FEC_{max} models, at a packet loss rate of 10%.

Figures A7 and A8 illustrate an increase in the FEC/FEC_{max} allocation to C1 and to C2. Allocation rates for C3 remain unchanged. The FEC/FEC_{max} allocation are based on the direct correlation scenario, as shown in Table A2. Note the large decrease in lower layer viewing and increase for most

schemes in the highest class, with reference to Figures A3 and A4, respectively, with a $Packet_{thres}$ of 500 bytes for all three classes using FEC_{max} allocation rates providing near complete C3 viewing.

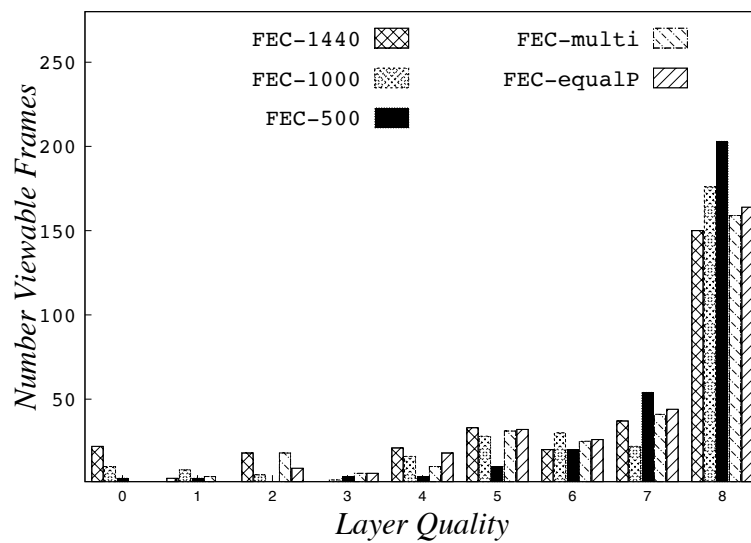


Figure A7. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{Packet_{thres}}$, at a packet loss rate of 10%. This image illustrates an increase in the FEC for C1 and C2 based on the worst case scenario, as shown in Table A2.

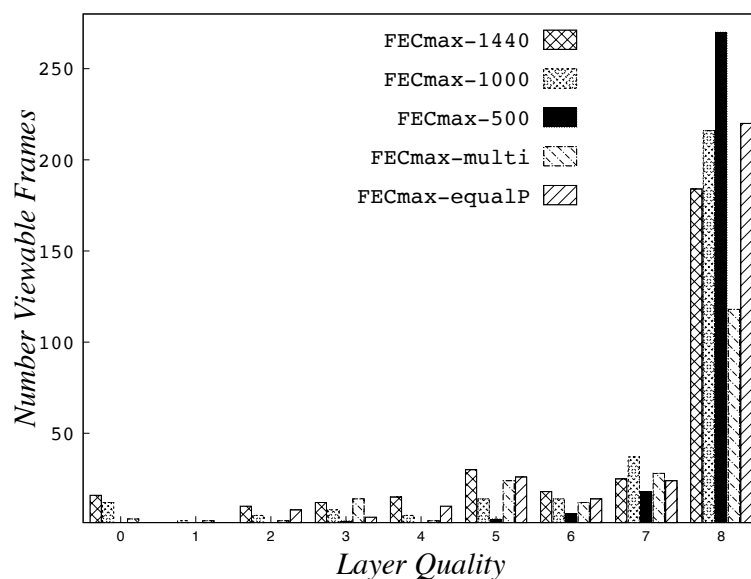


Figure A8. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{max}Packet_{thres}$, at a packet loss rate of 10%. This image illustrates an increase in the FEC_{max} for C1 and C2 based on the worst case scenario, as shown in Table A2.

Finally, Table A5 presents the cumulative transmission cost of the three classes based on the respective underlying FEC allocation rate, e.g., Table A1 (best case) and Table A2 (worst case). We note that, for FEC_{max} , Table A2 has increased the cumulative transmission cost of C1 by 32%, C2 by 19% and C3 by 8%. When we view these costs with respect to ALD from Table A3, we note a reduction in C1 for FEC_{max} of 72%, in C2 of 40% but an increase of 3% in C3. While the transmission cost for FEC_{max} in C1 with respect to SVC is an increase of 57%, when compared to the C1 cost for MDC (293%) and ALD (460%), this is a dramatic decrease in overall transmission cost for C1. Similar savings can be made in C2.

Table A5. Cumulative transmission cost of the three classes based on the respective underlying FEC allocation rate, e.g., Table A1 (best case) and Table A2 (worst case).

Table A1			Table A2		
		FEC	FEC _{max}	FEC	FEC _{max}
L8	Class 3	11,176,159	11,519,753	12,014,762	12,447,975
L7					
L6					
L5	Class 2	4,619,070	4,760,894	5,457,673	5,689,116
L4					
L3					
L2	Class 1	1,430,363	1,473,636	1,855,808	1,935,993
BL					

Appendix A.2. Varying Packet Size for a GOP of Eight

Table A6 is the GOP of eight equivalents of Table A4. Note the decrease in max loss rate per frame, as well as the decrease in overall transmission cost, as GOP increases. The remainder of this section compares GOP8 plots and figures to the previously shown GOP1 plots and figures.

Table A6. FEC and FEC_{max} models for all four *Packet_{thres}*, as well as *Packet_{equal}*. Values are illustrated for packets transmitted, additional packet header cost (60 bytes per packet) and individual packets sent, packets received, loss rate per class and maximum loss rate per frame for SVC-SC, for an overall packet loss rate of 10% and a GOP of eight.

SVC-SC 10% Different Packet Loss Rates Per Class, with a GOP of 8										
	FEC					FEC _{max}				
Byte Size	1440	1000	500	Multi	<i>Packet_{equal}</i>	1440	1000	500	Multi	<i>Packet_{equal}</i>
# C3 Trans Cost	4,679,425	4,679,944	4,681,985	4,679,896	4,680,918	4,824,226	4,824,655	4,826,978	4,824,952	4,825,324
# Packets	3303	4736	9418	4404	6021	3405	4880	9702	4534	6201
Header Cost	198,180	284,160	565,080	264,240	361,260	204,300	292,800	582,120	272,040	372,060
Packets Sent										
C1	392	561	1103	1103	2007	404	576	1132	1132	2067
C2	904	1294	2573	1294	2007	934	1335	2652	1335	2067
C3	2007	2881	5742	2007	2007	2067	2969	5918	2067	2067
Packets Received										
C1	345	495	984	975	1791	364	511	1020	1016	1858
C2	806	1163	2324	1153	1808	833	1203	2404	1201	1841
C3	1790	2559	5153	1821	1812	1836	2664	5285	1847	1876
Loss Rate (LR)										
C1	12%	12%	11%	12%	11%	10%	11%	10%	10%	10%
C2	11%	10%	10%	11%	10%	11%	10%	9%	10%	11%
C3	11%	11%	10%	9%	10%	11%	10%	11%	11%	09%
Max per Frame LR										
C1	50%	30%	28%	27%	22%	38%	31%	22%	22%	20%
C2	27%	18%	15%	20%	21%	22%	24%	18%	19%	18%
C3	21%	100%	15%	17%	22%	21%	18%	16%	25%	18%

Figures A9 and A10 illustrate a direct comparison to Figures A3 and A4. Note how a higher GOP value provides overall higher quantities of the higher layers, but still with excessive numbers of non-decodable frames. This is primarily due to the cascading effect of SD packetisation, where a higher packet loss rate, bursty loss, in one GOP will affect all frames for that GOP, thus mandating low or no decodable quality. Also note how the relatively small increase provided by FEC_{max} can show vastly improved higher layer viewing numbers, especially noticeable in Figure A10 for ‘500’ and ‘equalP’.

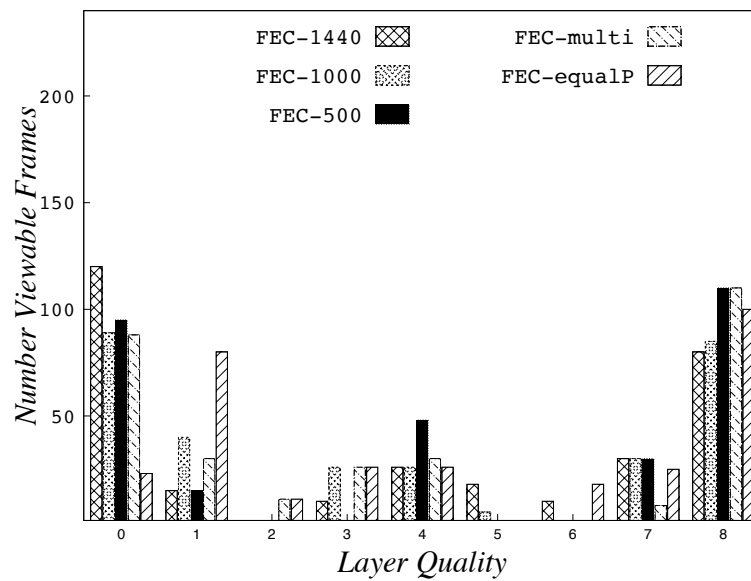


Figure A9. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{Packet_{thres}}$, at a packet loss rate of 10% and a GOP of 8.

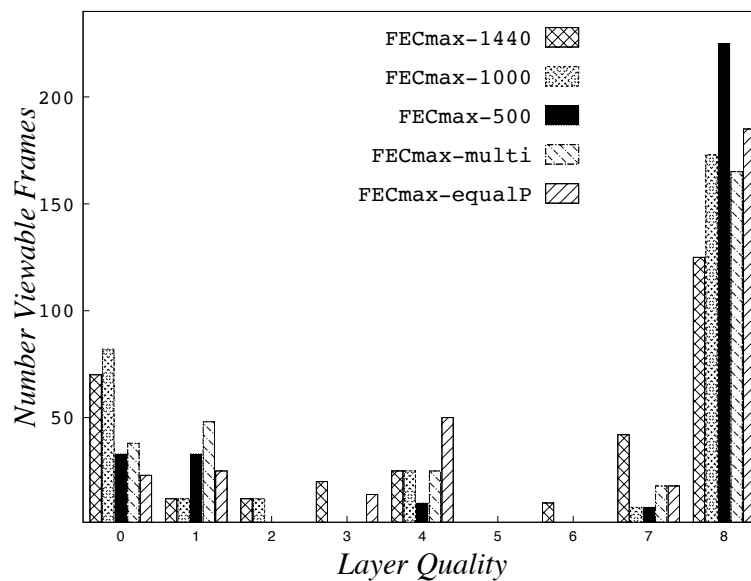


Figure A10. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{maxPacket_{thres}}$, at a packet loss rate of 10% and a GOP of 8.

Figures A11 and A12 illustrate a direct comparison to Figures A5 and A6. Note the consistency of layer quality that is provided by a larger GOP value (play out over all eight frames in the GOP), but with noticeably large variation in the quality. Again note the large increase in quality provided by the relatively small increase provided by FEC_{max} , especially for '500', 'multi' and 'equalP'. As previously mentioned, the cascading effect mandated by the frame interdependence within a GOP forces all frames within a GOP to the same layer value. This can be seen to be beneficial for high quality levels, but is detrimental to users when low quality is mandated.

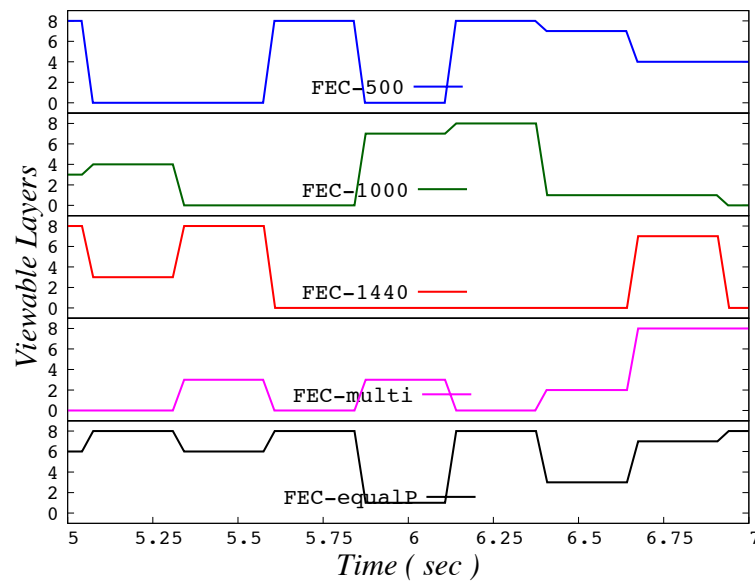


Figure A11. A two-second example of variation in viewable quality for all FEC schemes, at a packet loss rate of 10% and a GOP of 8.

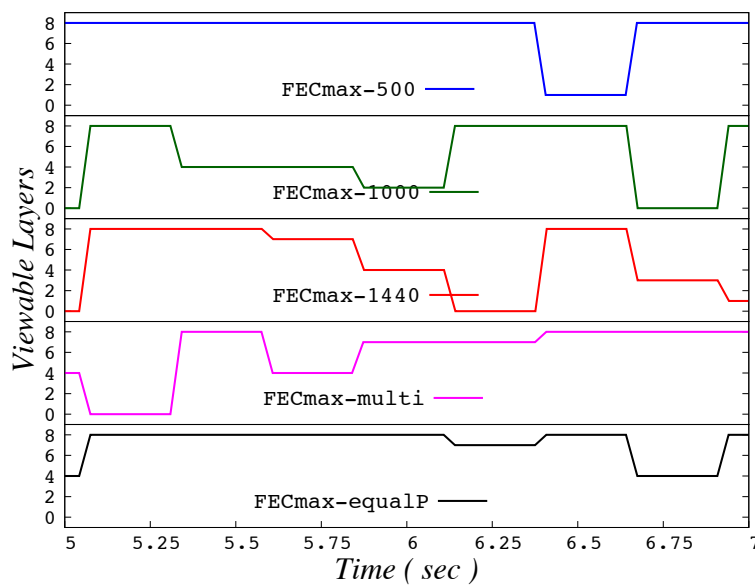


Figure A12. A two-second example of variation in viewable quality for all FEC_{max} schemes, at a packet loss rate of 10% and a GOP of 8.

Figures A13 and A14 illustrate a direct comparison to Figures A7 and A8, where the FEC levels have been increased in C1 and C2. For FEC_{max} , we increase the FEC in C1 and C2 to the maximum, while maintaining the standard level of FEC to C3, and note near continuous quality over the layers in C3 for all schemes except for $Packet_{thres}$ 1440 bytes. This illustrates that adaptation of the FEC levels in the lowest layers, or highest prioritised classes, is most beneficial to overall quality, and that lower levels of FEC can be utilised in the higher classes to increase, or maintain high levels, of viewable quality. Figure A15 illustrates the near consistency in viewable quality at quality layer 8, for all models with FEC_{max} . We begin now to see that adaptation in the GOP frame level, as well as the FEC level per class, can increase overall viewable quality.

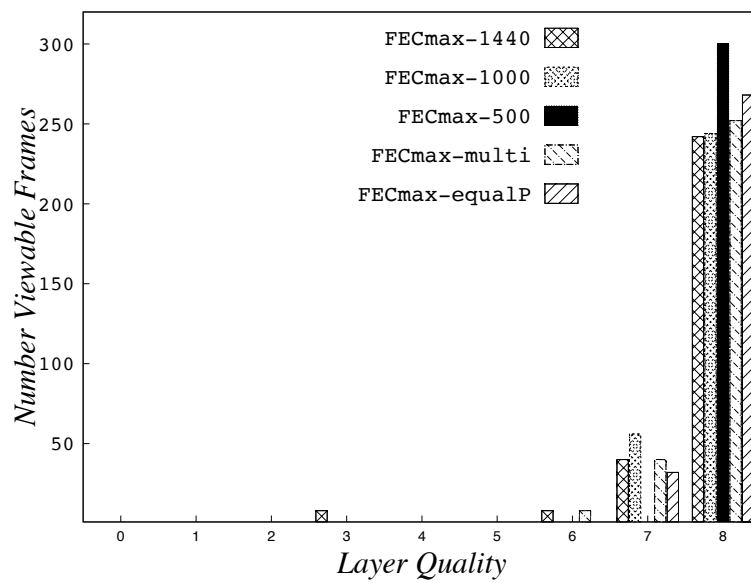


Figure A13. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{Packet_{thres}}$, at a packet loss rate of 10% and a GOP of 8. This image illustrates an increase in the FEC for C1 and C2 based on the worst case scenario, as shown in Table A2.

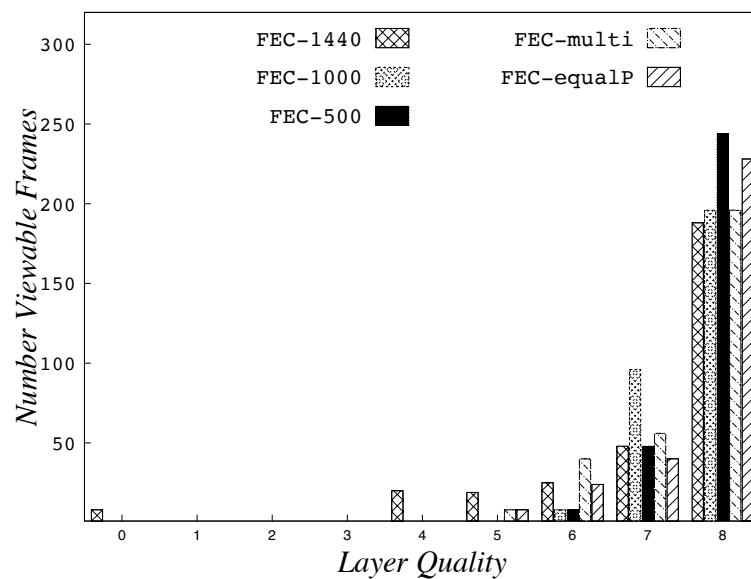


Figure A14. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{max}^{Packet_{thres}}$, at a packet loss rate of 10% and a GOP of 8. This image illustrates an increase in the FEC_{max} for C1 and C2 based on the worst case scenario, as shown in Table A2.

The results for GOP-8 illustrate that as GOP value increases, the same level of FEC provides for a noticeable increase in the viewable quality.

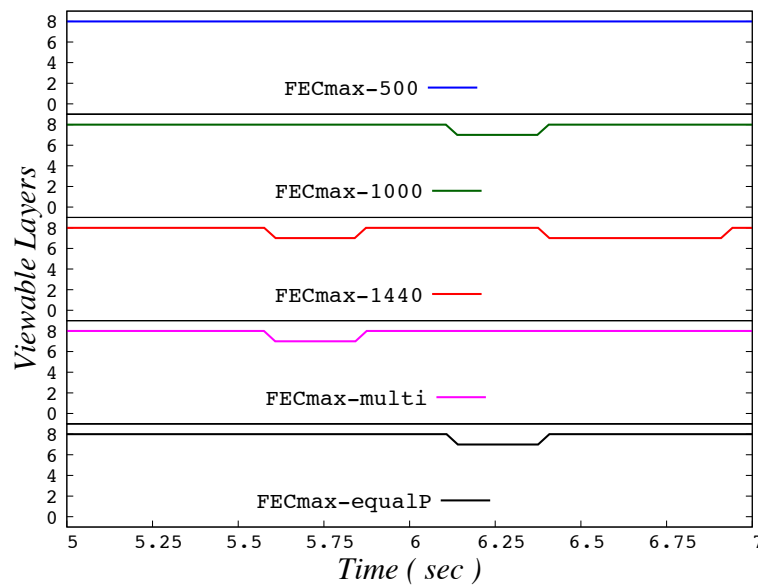


Figure A15. A two-second example of variation in viewable quality for all FEC_{max} models, at a packet loss rate of 10% and a GOP of 8. This illustrates an increase in the FEC_{max} for C1 and C2 based on the worst case scenario, as shown in Table A2.

Appendix A.3. Varying Packet Size for a GOP of Sixteen

Table A7 is the GOP of sixteen equivalent of Table A4. Note that, as we increase GOP value, there is a noticeable decrease in max loss rate per frame, as well as the decrease in overall transmission cost. The remainder of this section compares the GOP of 16 plots and figures to the previously shown GOP of 1 and GOP of 8 plots.

Table A7. FEC and FEC_{max} models for all four $Packet_{thres}$, as well as $Packet_{equal}$. Values are illustrated for packets transmitted, additional packet header cost (60 bytes per packet) and individual packets sent, packets received, loss rate per class and maximum loss rate per frame for SVC-SC, for an overall packet loss rate of 10% and a GOP of sixteen.

SVC-SC 10%—Different Packet Loss Rates Per Class, with a GOP of 16										
Byte Size	FEC					FEC _{max}				
	1440	1000	500	Multi	$Packet_{equal}$	1440	1000	500	Multi	$Packet_{equal}$
# C3 Trans Cost	3,720,325	3,720,920	3,722,521	3,720,844	3,721,817	3,835,615	3,835,926	3,837,584	3,836,188	3,836,895
# Packets	2611	3748	7471	3478	4794	2692	3862	7698	3583	4944
Header Cost	156,660	224,880	448,260	208,680	287,640	161,520	231,720	461,880	214,980	296,640
Packets Sent										
C1	307	438	869	869	1,598	317	452	893	893	1648
C2	706	1011	2013	1011	1598	727	1042	2075	1042	1648
C3	1598	2299	4589	1598	1598	1648	2368	4730	1648	1648
Packets Received										
C1	267	399	781	781	1445	287	404	804	787	1475
C2	636	901	1814	891	1432	658	931	1884	928	1471
C3	1420	2051	4123	1425	1426	1452	2119	4235	1480	1491
Loss Rate (LR)										
C1	13%	9%	10%	10%	10%	9%	11%	10%	12%	10%
C2	10%	11%	10%	12%	10%	9%	11%	9%	11%	11%
C3	11%	11%	10%	11%	11%	12%	11%	10%	10%	10%
Max per Frame LR										
C1	27%	31%	17%	21%	14%	27%	25%	21%	24%	17%
C2	27%	16%	16%	21%	15%	18%	16%	16%	18%	16%
C3	20%	16%	15%	15%	19%	20%	16%	14%	15%	20%

Figures A16 and A17 illustrate a direct comparison to Figures A9 and A10, based on the best case, lowest level, allocation of FEC as per Table A1. Note how high levels of layer 8 are available for FEC_{max} , even with the lowest level (best case) of FEC allocation. Also note how the loss is now being forced towards the highest layer in the lower classes, i.e., from layer 1, layer 4, such that the SD packetisation is now best equipped to deal with loss over a larger number of frames, as well as over a greater overall bitrate per GOP.

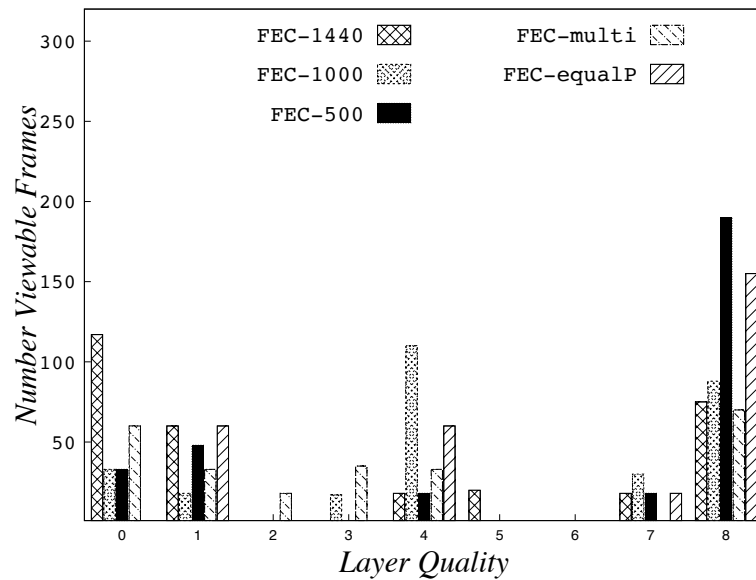


Figure A16. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{Packet_{thres}}$, at a packet loss rate of 10% and a GOP of 16.

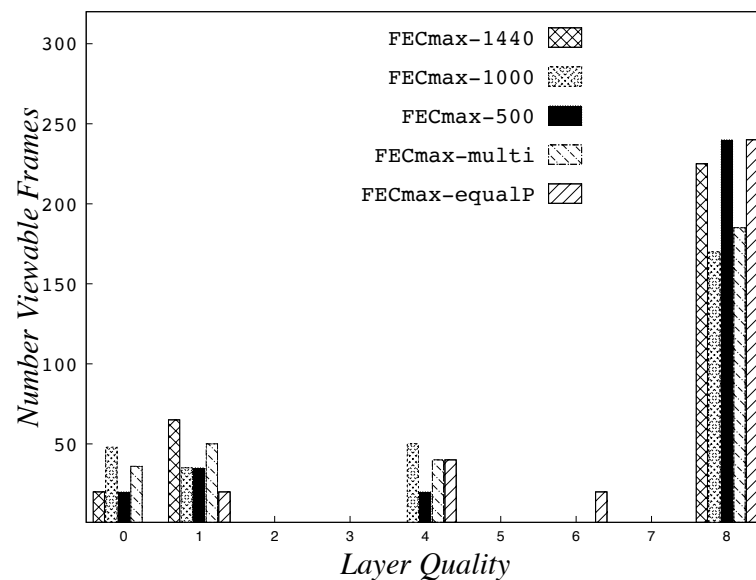


Figure A17. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{max}Packet_{thres}$, at a packet loss rate of 10% and a GOP of 16.

Figures A18 and A19 illustrate a two-second example of variation in viewable quality for all FEC schemes with minimum FEC allocation. Note how, for FEC_{max} , '500' and '1440' are now able to achieve consistency of quality at the highest layer for the duration of the time period shown, further

illustrating how a higher number of frames per GOP as well as an adaptive FEC allocation can provide consistency of quality over time.

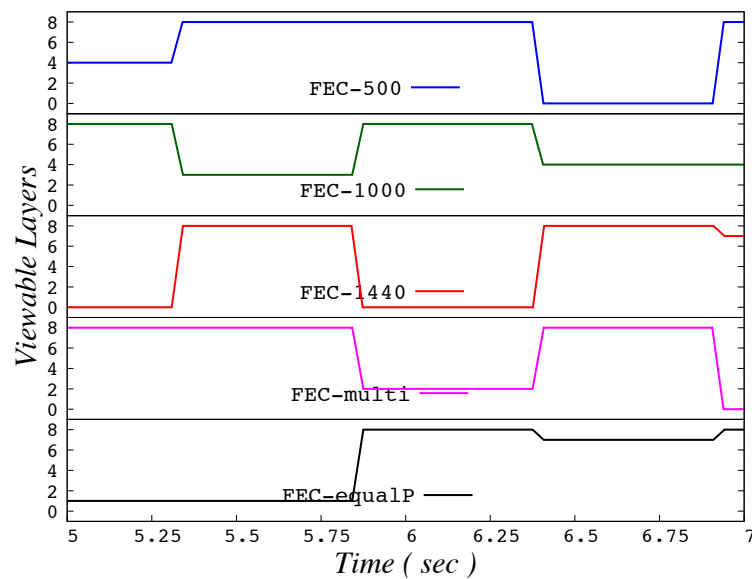


Figure A18. A two-second example of variation in viewable quality for all *FEC* schemes, at a packet loss rate of 10% and a GOP of 16.

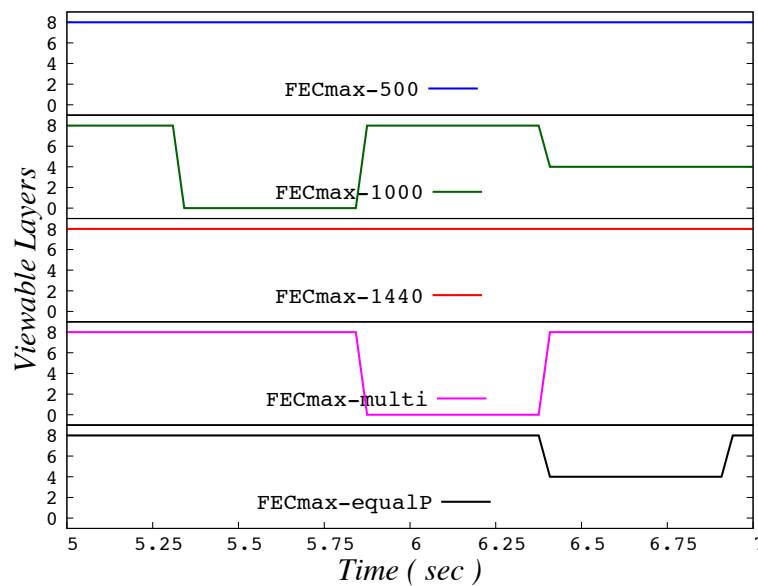


Figure A19. A two-second example of variation in viewable quality for all *FEC_{max}* schemes, at a packet loss rate of 10% and a GOP of 16.

Figures A20 and A21 illustrate an increase in the FEC in C1 and C2 to the maximum worst case level, while maintaining the standard level of FEC to C3 as per worst case allocation rates shown in Table A2. As can be seen in *FEC_{max}*, '500' and '1000' show no degradation in viewable quality, as all 300 frames can be shown in layer 8. This illustrates that, for this level of GOP, a lower level of FEC may be sufficient to mandate the nondegradation in viewable quality. It is also important to note that all schemes are now viewable within the highest class, i.e., within the highest resolution, thus variations in viewable quality are limited only to fidelity levels, thus further limiting the variation in noticeable

perceived quality to spatial rather than temporal issues. Figure A22 illustrates the near consistency in viewable quality for all models with FEC_{max} within the time period illustrated.

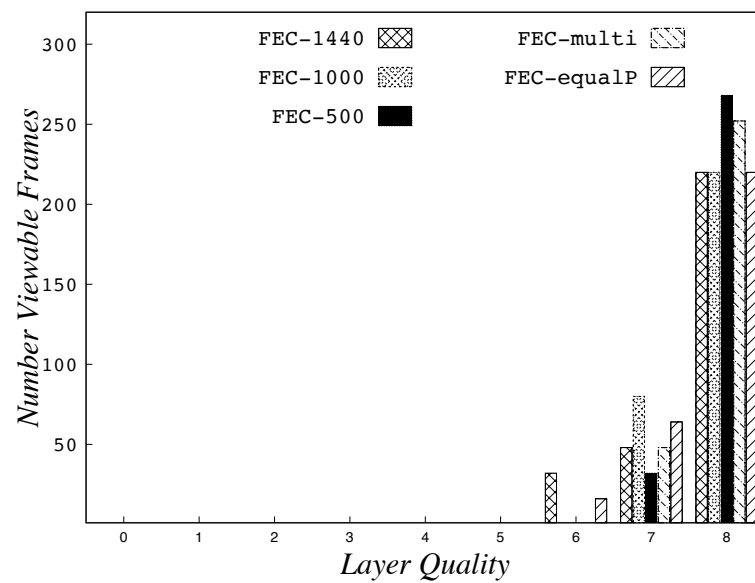


Figure A20. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{Packet_{thres}}$, at a packet loss rate of 10% and a GOP of 16. This image illustrates an increase in the FEC for C1 and C2 based on the worst case scenario, as shown in Table A2.

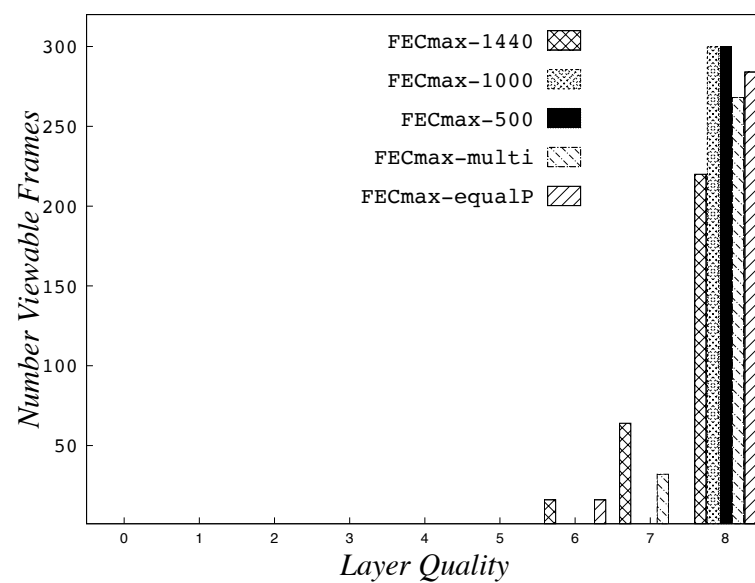


Figure A21. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{max}Packet_{thres}$, at a packet loss rate of 10% and a GOP of 16. This image illustrates an increase in the FEC_{max} for C1 and C2 based on the worst case scenario, as shown in Table A2.

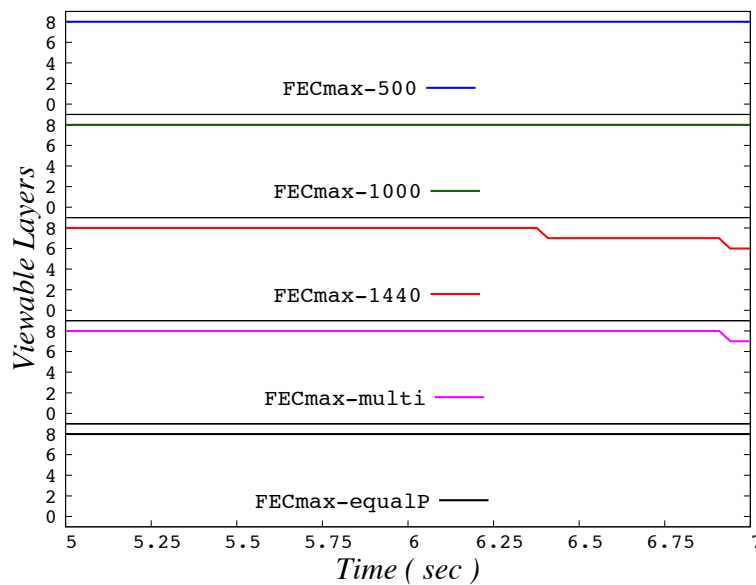


Figure A22. A two-second example of variation in viewable quality for all FEC_{max} models, at a packet loss rate of 10% and a GOP of 16. This illustrates an increase in the FEC_{max} for C1 and C2 based on the worst case scenario, as shown in Table A2.

In Figure A21, we have reached an error resiliency threshold for FEC_{max} models 1000 bytes, 500 bytes using worst case FEC allocation, as we have attained continuous maximum quality decoding of layer 8. In Figure A20, we also note that the lower FEC allocation rate is sufficient to provide full C1 and C2 decoding, while only C3 requires the additional FEC_{max} allocation to maximise quality. Thus, we now have an adaptive mechanism by which to maximise quality for the individual stream classes. As we have seen before, as we increase the GOP value, the same levels of FEC allocation provide an increase in viewable quality, as loss can now be distributed over more frames, or specifically over more bytes, thus forcing the loss to the FEC allocated levels rather than to specific frames or layers within the GOP.

Appendix A.4. Varying Packet Size for a GOP of 32

Table A8 contains the details for a GOP of 32. Note the decrease in max loss rate per frame, as well as the decrease in overall transmission cost, as GOP increases.

Figures A23 and A24 illustrate an increase in the FEC in C1 and C2 to the maximum worst case level, while maintaining the standard level of FEC to C3. Figure A25 illustrates two seconds of consistency in viewable quality for all models with FEC_{max} . As seen in Figure A24, only $Packet_{thres}$ 1440 was unable to provide continuous layer 8 (for one GOP layer 7 was decodable). We further evaluated GOP of 32 (not shown) and found that when we increased the FEC_{max} allocation of C1 from 17% and 22% to 22% and 27% respectively, based on Table A1, that excluding $Packet_{thres}$ 1440, all the other models were able to decode all but one of the GOP at layer 8. Further increasing C1 from 22% and 27% to 27% and 33%, respectively, and increasing C2 from 17%, 22% and 27% to 22%, 27% and 33%, respectively, increased viewable quality of the three models to full layer 8 decoding, while $Packet_{thres}$ 1440 bytes and 'multi' were able to decode all but one of the GOP at layer 8. This further illustrates the adaptive FEC allocation, as well as the packetisation byte-allocation, and mechanisms proposed can find the optimal level of FEC, but, as defined, a non-default initial level is required prior to GOP value determination. It can be concluded that higher GOP values will reduce overall FEC levels even further.

Table A8. FEC and FEC_{max} models for all four $Packet_{thres}$, as well as $Packet_{equal}$. Values are illustrated for packets transmitted, additional packet header cost (60 bytes per packet) and individual packets sent, packets received, loss rate per class and maximum loss rate per frame for SVC-SC, for an overall packet loss rate of 10% and a GOP of 32.

SVC-SC 10%—Different Packet Loss Rates Per Class, with a GOP of 32										
Byte Size	FEC					FEC_{max}				
	1440	1000	500	Multi	$Packet_{equal}$	1440	1000	500	Multi	$Packet_{equal}$
# C3 Trans Cost	3,158,123	3,158,382	3,160,689	3,158,353	3,159,005	3,255,771	3,256,625	3,258,020	3,256,052	3,256,772
# Packets	2210	3174	6331	2936	4098	2276	3270	6524	3022	4218
Header Cost	132,600	190,440	379,860	176,160	245,880	136,560	196,200	391,440	181,320	253,080
Packets Sent										
C1	255	364	723	723	1366	263	373	742	742	1406
C2	589	847	1689	847	1366	607	874	1742	874	1406
C3	1366	1963	3919	1366	1366	1406	2023	4040	1406	1406
Packets Received										
C1	219	319	646	636	1218	230	334	673	652	1269
C2	522	771	1524	758	1215	546	784	1568	786	1252
C3	1231	1736	3519	1223	1233	1255	1796	3623	1257	1255
Loss Rate (LR)										
C1	14%	12%	11%	12%	11%	13%	10%	9%	12%	10%
C2	11%	9%	10%	11%	11%	10%	10%	10%	10%	11%
C3	10%	12%	10%	10%	10%	11%	11%	10%	11%	11%
Max per Frame LR										
C1	36%	23%	16%	19%	15%	22%	27%	14%	18%	15%
C2	21%	13%	14%	16%	16%	16%	14%	17%	15%	16%
C3	14%	16%	12%	14%	14%	16%	14%	12%	15%	14%

Rather than illustrate results for both FEC allocation tables, Tables A1 and A2, we will just illustrate the results for Table A2. We note that, for the results based on Table A1, only one specific GOP was unable to provide layer 8 decoding for FEC_{max} $Packet_{thres}$ 1000 bytes and 500 bytes.

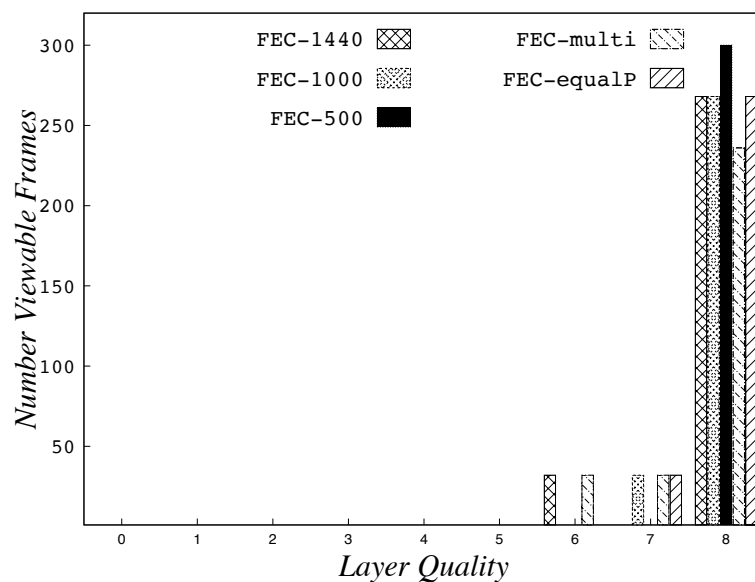


Figure A23. Example of the number of viewable layers for SVC-SC, for each of the FEC_{thres} , at a packet loss rate of 10% and a GOP of 32. This image illustrates an increase in the FEC for C1 and C2 based on the worst case scenario, as shown in Table A2.

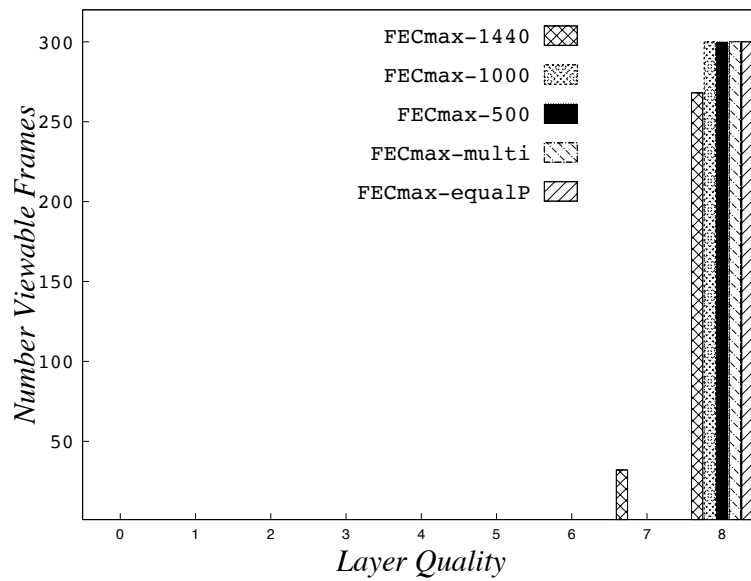


Figure A24. Example of the number of viewable layers for SVC-SC, for each of the $FEC_{max}Packet_{thres}$, at a packet loss rate of 10% and a GOP of 32. This image illustrates an increase in the FEC_{max} for C1 and C2 based on the worst case scenario, as shown in Table A2.

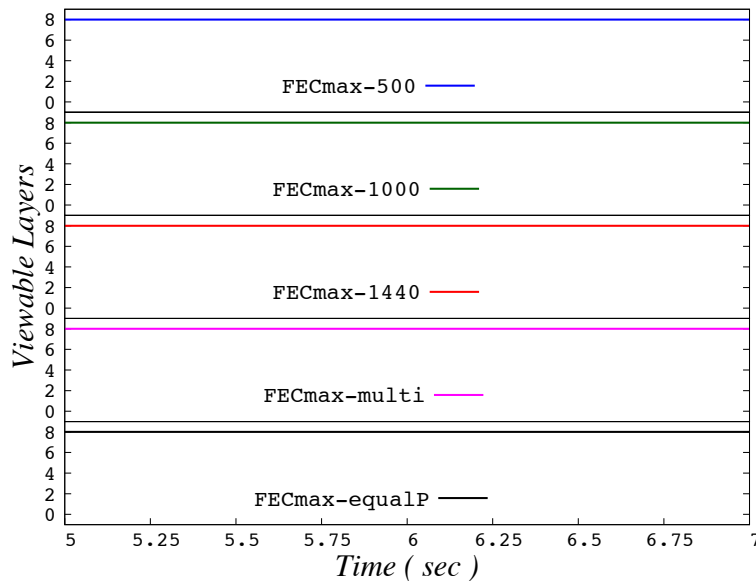


Figure A25. A two-second example of variation in viewable quality for all FEC_{max} models, at a packet loss rate of 10% and a GOP of 32. This illustrates an increase in the FEC_{max} for C1 and C2 based on the worst case scenario, as shown in Table A2.

Appendix A.5. Appendix Conclusions

We noted, with a GOP of 1, that the worst case FEC allocation could not provide continuous layer 8 quality, while, with a GOP of 8, the same level of FEC allocation provides near continuous layer 8 quality for all schemes. For a GOP of 16 and GOP of 32, we were able to reduce FEC allocation rates for C1 (GOP of 16) and for C1 and C2 (GOP of 32) while mandating maximised viewable quality, thus illustrating both adaptation in FEC allocation and packet byte-size allocation by which viewable quality can be preserved and stabilised during moment of packet loss.

As GOP increases, we note that the range of maximum per frame loss rates is smaller, thus leveraging the benefits of sharing the GOP loss rate over more frames. Note that the maximum

loss rate per frame is actually the maximum loss rate per GOP, but, because of how we packetise the layer and frames per GOP, both per frame and per GOP values are equal. In these results, we did not increase the FEC or FEC_{max} for C3. Increasing error resiliency in C3 will only increase viewable quality in C3, while increases in error resiliency in lower classes benefit both the individual lower class and all higher classes. As seen, as GOP values increases, the minimum level of FEC or FEC_{max} for C3 was adequate for most models to stream at maximum quality.

Finally, we note that, for a GOP of 1, a transmission cost of 12,447,975 bytes is required for FEC_{max} , while a GOP of 32 will mandate a reduced transmission cost of 3,494,525 bytes, thus illustrating that lower FEC rates per layer and a lower overall transmission cost can provide maximised viewable quality at higher GOP values.

References

1. Wu, D.; Hou, Y.; Zhu, W.; Zhang, Y.; Peha, J. Streaming Video Over the Internet: Approaches and Directions. *IEEE Trans. Circuits Syst. Video Technol.* **2001**, *11*, 282–300.
2. Wiegand, T.; Sullivan, G. The H.264/AVC Video Coding Standard [Standards in a Nutshell]. *IEEE Signal Process. Mag.* **2007**, *24*, 148–153.
3. Advanced Video Coding for Generic Audiovisual Services. Available online: <http://www.itu.int/rec/T-REC-H.264> (accessed on 18 January 2018).
4. Marpe, D.; Wiegand, T.; Sullivan, G. The H.264/MPEG4 Advanced Video Coding Standard and its Applications. *Commun. Mag.* **2006**, *44*, 134–143.
5. Sánchez de la Fuente, Y.; Schierl, T.; Hellge, C.; Wiegand, T.; Hong, D.; De Vleeschauwer, D.; Van Leekwijck, W.; Le Louédec, Y. iDASH: Improved Dynamic Adaptive Streaming over HTTP using Scalable Video Coding. In Proceedings of the Second Annual ACM Conference on Multimedia Systems, Santa Clara, CA, USA, 23–25 February 2011; pp. 257–264.
6. Lekharu, A.; Kumar, S.; Sur, A.; Sarkar, A. A QoE Aware SVC Based Client-side Video Adaptation Algorithm for Cellular Networks. In Proceedings of the 19th International Conference on Distributed Computing and Networking, Los Angeles, CA, USA, 4–7 January 2018; pp. 27:1–27:4.
7. Radha, H.; Chen, Y.; Parthasarathy, K.; Cohen, R. Scalable Internet Video using MPEG-4. *Signal Process.* **1999**, *15*, 95–126.
8. ITU-T SVC Standard. Available online: <https://www.itu.int/ITU-T/recommendations/index.aspx?ser=H> (accessed on 18 January 2018).
9. Schwarz, H.; Marpe, D.; Wiegand, T. Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. *IEEE Trans. Circuits Syst. Video Technol.* **2007**, *17*, 1103–1120.
10. Wang, Y.K.; Hannuksela, M.; Pateux, S.; Eleftheriadis, A.; Wenger, S. System and Transport Interface of SVC. *IEEE Trans. Circuits Syst. Video Technol.* **2007**, *17*, 1149–1163.
11. Draft RTP Payload Format for SVC. Available online: <http://tools.ietf.org/html/draft-ietf-avt-rtp-svc-18#page-6> (accessed on 18 January 2018).
12. Wenger, S.; Wang, Y.K.; Hannuksela, M. RTP Payload Format for H.264/SVC Scalable Video Coding. *J. Zhejiang Univ. Sci. A* **2006**, *7*, 657–667.
13. Schierl, T.; Hellge, C.; Mirta, S.; Gruneberg, K.; Wiegand, T. Using H.264/AVC-based Scalable Video Coding (SVC) for Real Time Streaming in Wireless IP Networks. In Proceedings of the 2007 IEEE International Symposium on Circuits and Systems, New Orleans, LA, USA, 27–30 May 2007; pp. 3455–3458.
14. Helle, P.; Lakshman, H.; Siekmann, M.; Stegemann, J.; Hinz, T.; Schwarz, H.; Marpe, D.; Wiegand, T. A Scalable Video Coding Extension of HEVC. In Proceedings of the Data Compression Conference (DCC), Snowbird, UT, USA, 20–22 March 2013; pp. 201–210.
15. Nightingale, J.; Wang, Q.; Grecos, C. Scalable HEVC (SHVC)-Based Video Stream Adaptation in Wireless Networks. In Proceedings of the 2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), London, UK, 8–11 September 2013; pp. 3573–3577.
16. Kessentini, A.; Damak, T.; Ayed, M.A.B.; Masmoudi, N. Scalable high efficiency video coding (SHEVC) performance evaluation. In Proceedings of the 2015 World Congress on Information Technology and Computer Applications (WCITCA), Hammamet, Tunisia, 11–13 June 2015; pp. 1–4.

17. Boyce, J.M.; Ye, Y.; Chen, J.; Ramasubramonian, A.K. Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 20–34.
18. Puri, R.; Ramchandran, K.; Lee, K. Forward Error Correction (FEC) Codes Based Multiple Description Coding for Internet Video Streaming and Multicast. *Signal Process.* **2001**, *16*, 745–762.
19. Wang, B.; Kurose, J.; Shenoy, P.; Towsley, D. Multimedia Streaming via TCP: An Analytic Performance Study. *ACM Trans. Multimed. Comput. Commun. Appl.* **2008**, *4*, 16.
20. Kuschnig, R.; Kofler, I.; Hellwagner, H. Evaluation of HTTP-based Request-response Streams for Internet Video Streaming. In Proceedings of the ACM Multimedia Systems Conference (MMSys'11), San Jose, CA, USA, 23–25 February 2011; pp. 245–256.
21. Przybylski, M.; Belter, B.; Binczewski, A. Shall We Worry about Packet Reordering? *Comput. Methods Sci. Technol* **2005**, *11*, 141–146.
22. Loguinov, D.; Radha, H. End-to-End Internet Video Traffic Dynamics: Statistical Study and Analysis. In Proceedings of the IEEE Information Communications Conference (INFOCOM 2002), New York, NY, USA, 23–27 June 2002; pp. 723–732.
23. Chen, W.; Mitra, U.; Neely, M. Packet Dropping Algorithms for Energy Savings. In Proceedings of the 2006 IEEE International Symposium on Information Theory, Seattle, WA, USA, 9–14 July 2006; pp. 227–231.
24. Xiao, S.; Wang, H.; Kuo, C.C. Priority Ordering and Packetization for Scalable Video Multicast with Network Coding. In *Advances in Multimedia Information Processing*; Ip, H.H.S., Au, O.C., Leung, H., Sun, M.T., Ma, W.Y., Hu, S.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 520–529.
25. Wenger, S.; Wang, Y.; Schierl, T. Transport and Signaling of SVC in IP Networks. *IEEE Trans. Circuits Syst. Video Technol.* **2007**, *17*, 1164–1173.
26. Trestian, R.; Muntean, G.; Ormond, O. Signal Strength-based Adaptive Multimedia Delivery Mechanism. In Proceedings of the IEEE 34th Conference on Local Computer Networks, Zurich, Switzerland, 20–23 October 2009; pp. 297–300.
27. Nunes, B.; Mendonca, M.; Xuan-Nam, N.; Obraczka, K.; Turletti, T. A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 1617–1634.
28. Karakus, M.; Durresi, A. Quality of Service (QoS) in Software Defined Networking (SDN): A survey. *J. Netw. Comput. Appl.* **2017**, *80*, 200–218.
29. Quang, P.T.A.; Piamrat, K.; Singh, K.D.; Viho, C. Video Streaming Over Ad Hoc Networks: A QoE-Based Optimal Routing Solution. *IEEE Trans. Veh. Technol.* **2017**, *66*, 1533–1546.
30. Sen, S.; Gilani, S.; Srinath, S.; Schmitt, S.; Banerjee, S. Design and Implementation of an “Approximate” Communication System for Wireless Media Applications. In Proceedings of the ACM SIGCOMM 2010 Conference, New Delhi, India, 30 August–3 September 2010; pp. 15–26.
31. Jamieson, K.; Balakrishnan, H. PPR: Partial Packet Recovery for Wireless Networks. In Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, Kyoto, Japan, 27–31 August 2007; pp. 409–420.
32. Nafaa, A.; Taleb, T.; Murphy, L. Forward Error Correction Strategies for Media Streaming over Wireless Networks. *IEEE Commun. Mag.* **2008**, *46*, 72–79.
33. Goyal, V. Multiple Description Coding: Compression Meets the Network. *Signal Process. Mag.* **2002**, *18*, 74–93.
34. Wang, Y.; Reibman, A.; Lin, S. Multiple Description Coding for Video Delivery. *Proc. IEEE* **2005**, *93*, 57–70.
35. Kazemi, M.; Shirmohammadi, S.; Sadeghi, K. A Review of Multiple Description Coding Techniques for Error-resilient Video Delivery. *Multimed. Syst.* **2014**, *20*, 283–309.
36. Ghahremani, S.; Ghanbari, M. Error resilient video transmission in ad hoc networks using layered and multiple description coding. *Multimed. Tools Appl.* **2017**, *76*, 9033–9049.
37. Heide, J.; Sorensen, J.; Krigslund, R.; Popovski, P.; Larsen, T.; Chakareski, J. Cooperative Media Streaming using Adaptive Network Compression. In Proceedings of the 2008 International Symposium World of Wireless, Mobile and Multimedia Networks, Newport Beach, CA, USA, 23–26 June 2008; pp. 1–7.
38. Bai, H.; Wang, A.; Zhao, Y.; Pan, J.S.; Abraham, A. *Distributed Multiple Description Coding; Principles, Algorithms and Systems*; Springer: New York, NY, USA 2011.
39. Hussain, M.; Hameed, A. Adaptive video-aware forward error correction code allocation for reliable video transmission. *Signal Image Video Process.* **2018**, *12*, 161–169.
40. Kirihaara, K.; Masuyama, H.; Kasahara, S.; Takahashi, Y. FEC Recovery Performance for Video Streaming Services Based on H. 264/SVC. In *Recent Advances on Video Coding*; InTech: Kyoto, Japan, 2011.

41. Sorensen, J.; Ostergaard, J.; Popovski, P.; Chakareski, J. Multiple Description Coding with Feedback Based Network Compression. In Proceedings of the Global Telecommunications Conference (GLOBECOM 2010), Miami, FL, USA, 6–10 December 2010; pp. 1–6.
42. Majid, M.; Owais, M.; Anwar, S.M. Visual saliency based redundancy allocation in HEVC compatible multiple description video coding. In *Multimedia Tools and Applications*; Springer: Berlin, Germany, 2017.
43. Lin, C.; Zhao, Y.; Xiao, J.; Tillo, T. Region-based multiple description coding for multiview video plus depth video. *IEEE Trans. Multimed.* **2017**, *PP*, 1.
44. Puri, R. Multiple Description Source Coding using Forward Error Correction Codes. In Proceedings of the Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 24–27 October 1999; pp. 342–346.
45. Li, Z.; Khisti, A.; Girod, B. Forward Error Protection for Low-delay Packet Video. In Proceedings of the 2010 18th International Packet Video Workshop (PV), Hong Kong, China, 13–14 December 2010.
46. Reed, I.; Solomon, G. Polynomial Codes Over Certain Finite Fields. *J. Soc. Ind. Appl. Math.* **1960**, *8*, 300–304.
47. Shokrollahi, A. Raptor Codes. *IEEE Trans. Inf. Theory* **2006**, *52*, 2551–2567.
48. Chen, S.T.; Chiao, H.T.; Chang, S.Y.; Sun, H.M. Fast and low-complexity encoding of Raptor codes based on operation lists for known source block lengths. *Multimed. Tools Appl.* **2017**, *76*, 9219–9243.
49. Park, K.; Wang, W. AFEC: An Adaptive Forward Error Correction Protocol for End-to-end Transport of Real-time Traffic. In Proceedings of the 7th International Conference on Computer Communications and Networks, Lafayette, LA, USA, 12–15 October 1998; pp. 196–205.
50. Li, L.; Han, Q.; Niu, X. Enhanced Adaptive FEC Based Multiple Description Coding for Internet Video Streaming over Wireless Network. In Proceedings of the 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), Darmstadt, Germany, 15–17 October 2010; pp. 478–481.
51. Chou, P.; Wang, H.; Padmanabhan, V.N. Layered Multiple Description Coding. In Proceedings of the IEEE International Packet Video Workshop (PV2003), Nantes, France, 28–29 April 2003.
52. Albanese, A.; Blomer, J.; Edmonds, J.; Luby, M.; Sudan, M. Priority Encoding Transmission. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, New Mexico, 20–22 November 1994; pp. 604–612.
53. Zhao, Z.; Ostermann, J.; Chen, H. Multiple Description Scalable Coding for Video Streaming. In Proceedings of the 2009 10th Workshop Image Analysis for Multimedia Interactive Services, London, UK, 6–8 May 2009; pp. 21–24.
54. Zhao, Z.; Ostermann, J. Video Streaming Using Standard-compatible Scalable Multiple Description Coding based on SVC. In Proceedings of the 17th IEEE International Conference on Image Processing (ICIP), Hong Kong, China, 26–29 September 2010; pp. 1293–1296.
55. Berkin Abanoz, T.; Murat Tekalp, A. SVC-based Scalable Multiple Description Video Coding and Optimization of Encoding Configuration. *Signal Process.* **2009**, *24*, 691–701.
56. Quinlan, J.; Zahran, A.; Sreenan, C. ALD: Adaptive Layer Distribution for Scalable Video. In Proceedings of the ACM Multimedia Systems Conference (MMSys'13), Oslo, Norway, 27 February–1 March 2013; pp. 202–213.
57. Gan, T.; Gan, L.; Ma, K. Reducing Video-quality Fluctuations for Streaming Scalable Video using Unequal Error Protection, Retransmission, and Interleaving. *IEEE Trans. Image Process.* **2006**, *15*, 819–832.
58. Quinlan, J.; Zahran, A.; Sreenan, C. SDC: Scalable Description Coding for Adaptive Streaming Media. In Proceedings of the 19th IEEE International Packet Video Workshop (PV2012), Munich-Garching, Germany, 10–11 May 2012; pp. 59–64.
59. Quinlan, J.; Zahran, A.; Sreenan, C. ALD: Adaptive Layer Distribution for Scalable Video. In *ACM Multimedia Systems*; Springer: Berlin, Germany, **2014**, pp. 1–20.
60. Steinbach, E.; Färber, N.; Girod, B. Adaptive Payout for Low Latency Video Streaming. In Proceedings of the 2001 International Conference on Image Processing, Thessaloniki, Greece, 7–10 October 2001; pp. 962–965.
61. Sreenan, C.; Chen, J.C.; Agrawal, P.; Narendran, B. Delay Reduction Techniques for Payout Buffering. *IEEE Trans. Multimed.* **2000**, *2*, 88–100.
62. Dai, M.; Loguinov, D.; Radha, H. Rate-Distortion Analysis and Quality Control in Scalable Internet Streaming. *IEEE Trans. Multimed.* **2006**, *8*, 1135–1146.

63. Wang, Y.; Lin, T.L.; Cosman, P. Packet Dropping for H.264 Videos Considering both Coding and Packet-loss Artifacts. In Proceedings of the 2010 18th International Packet Video Workshop (PV), Hong Kong, China, 13–14 December 2010; pp. 165–172.
64. Helix Mobile Server. Available online: <https://www.reálnetworks.com/products-services/helix/eulas> (accessed on 12 January 2018).
65. Olshefski, D.; Nieh, J.; Agrawal, D. Inferring Client Response Time at the Web Server. In Proceedings of the 2002 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, Marina Del Rey, CA, USA, 15–19 June 2002; pp. 160–171.
66. IBM Unix Network Performance Analysis. Available online: <https://www.ibm.com/developerworks/aix/library/au-networkperfanalysis/> (accessed on 12 January 2018).
67. Muntean, G. Effect of Delivery Latency, Feedback Frequency and Network Load on Adaptive Multimedia Streaming. In Proceedings of the 32nd IEEE Conference on Local Computer Networks, Dublin, Ireland, 15–18 October 2007; pp. 421–427.
68. Basso, S.; Meo, M.; De Martin, J. Strengthening Measurements from the Edges: Application-level Packet Loss Rate Estimation. *SIGCOMM Comput. Commun. Rev.* **2013**, *43*, 45–51.
69. Zhang, Y.; Mao, Z.M.; Zhang, M. Detecting Traffic Differentiation in Backbone ISPs with NetPolice. In Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference, Chicago, IL, USA, 4–6 November 2009; pp. 103–115.
70. Madhyastha, H.; Isdal, T.; Piatek, M.; Dixon, C.; Anderson, T.; Krishnamurthy, A.; Venkataramani, A. iPlane: An Information Plane for Distributed Services. In Proceedings of the 7th Symposium on Operating Systems Design and Implementation, Berkeley, CA, USA, 6–8 November 2006; pp. 367–380.
71. Andersen, D.; Balakrishnan, H.; Kaashoek, F.; Morris, R. Resilient Overlay Networks. In Proceedings of the Eighteenth ACM Symposium on Operating Systems Principles, New York, NY, USA, 21–24 October 2001; pp. 131–145.
72. Wang, Y.; Huang, C.; Li, J.; Ross, K. Queen: Estimating Packet Loss Rate between Arbitrary Internet Hosts. In *Passive and Active Network Measurement*; Moon, S., Teixeira, R., Uhlig, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 57–66.
73. Chiang, Y.R.; Ke, C.H.; Yu, Y.S.; Chen, Y.S.; Pan, C.J. A multipath transmission scheme for the improvement of throughput over SDN. In Proceedings of the 2017 International Conference on Applied System Innovation (ICASI), Sapporo, Japan, 13–17 May 2017; pp. 1247–1250.
74. Malik, A.; Ahmed, J.; Qadir, J.; Ilyas, M.U. A measurement study of open source SDN layers in OpenStack under network perturbation. *Comput. Commun.* **2017**, *102*, 139–149.
75. Video Traces for Network Performance Evaluation. Available online: <ftp://ftp.tnt.uni-hannover.de/pub/svc/testsequences/> (accessed on 11 May 2017).
76. Rejaie, R.; Handley, M.; Estrin, D. RAP: An end-to-end rate-based congestion control mechanism for realtime streams in the Internet. In Proceedings of the Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies, New York, NY, USA, 21–25 March 1999; pp. 1337–1345.
77. Reichel, J.; Schwarz, H.; Wien, M. *Joint Scalable Video Model 9 (JSVM 9_19_15)*; Joint Video Team: Klagenfurt, Austria, 2007.
78. The Network Simulator-ns-2. Available online: http://nslam.sourceforge.net/wiki/index.php/Main_Page (accessed on 14 August 2017).
79. Ke, C. myEvalSVC: An Integrated Simulation Framework for Evaluation of H.264/SVC Transmission. *KSII Trans. Internet Inf. Syst.* **2012**, *6*, 379–394.
80. Salomon, D. *Guide to Data Compression Methods*; Springer: Berlin, Germany, 2002.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).